April 25, 1999

# Proposal for U.S.-ATLAS Computing

**Prepared by**
**T. LeCompte, D.M. Malon, E.N. May, L.E Price, R. Wagner**
*Argonne National Laboratory, Argonne, IL*
**P. Calafiura, M.G. Gilchriese, I. Hinchliffe, C. Leggate,**
S. Loken, J.L. Siegrist, M.D. Shapiro, G.H. Trilling, C.E. Tull
*Lawrence Berkeley National Laboratory, Berkeley, CA*
**S. Adler, S. Rajagopolan**
*Brookhaven National Laboratory, Upton, NY*
**F. Merritt, J. Pilcher**
*University of Chicago, Chicago, IL*
**S. McKee**
*University of Michigan, Ann Arbor, MI*
**K. Sliwa**
*Tufts U., Boston MA.*

## Contents

# 1 Executive Summary

We present here a proposal from the U.S.-ATLAS collaboration for computing. Our first goal is to ensure that U.S.-ATLAS physicists be able to participate fully in data analysis and physics discoveries at the LHC. Secondly, we intend that U.S.-ATLAS should contribute as strongly to computing as to major ATLAS detector subsystems. The combination of experience and expertise of the U.S. groups, based on other high energy and nuclear physics experiments, provides the opportunity for U.S.-ATLAS to make major contributions to ATLAS computing.

The principal elements of our proposal are the software development effort in the U.S. and the data storage and the computing infrastructure required to support that software development as well as to provide for the needs of U.S.-ATLAS to store and analyze data.

Our funding request is dominated by the need for professional software and computing personnel. We have estimated the appropriate scope of work based on experience with current experiments (BaBar, CDF/D0, STAR) and extrapolation to ATLAS needs. Both a top-down approach, noting that the U.S. is about 20% of the total ATLAS effort, and a bottom-up approach, based on proposed tasks, set the scale of the proposed U.S. software development effort at 120 man-years of physicist effort, and 80 man-years of professional software developer effort, spread over the next six years. Early adoption of full Object-Oriented (OO) techniques by U.S.-ATLAS physicists is a critical element of our plan. This effort will take maximum advantage of the considerable U.S. computing investment in BaBar, CDF/D0, and STAR. U.S.-ATLAS expects to recruit experienced software professionals from these efforts, as they become available, and has tailored its plans with this idea in mind.

Negotiations with ATLAS management on core software deliverables have been initiated. Our provisional plan is to take major responsibilities in the areas of Control/Framework and Database.

We have also chosen a short-term software pilot project (data access, analysis and simulation of Tile Calorimeter test-beam data) to encourage U.S. physicists to concentrate their work on fully OO code. This project can expand to accommodate additional test-beam analyses in the future.

An aggressive program of OO training will be implemented to ensure full involvement of U.S. physicists in the software development.

The hardware infrastructure needed over the next three years is modest. It could be provided at little additional cost by leveraging resources already available. For example, LBNL's Parallel Distributed Systems Facility (PDSF) facility appears to be of the proper size to meet our immediate needs. This hardware and the associated operations and support professionals will form the core of our near-term efforts in code maintenance and distribution, database deployment, framework development, physicist training and user support. In the longer term (beyond the next three years), the U.S.-ATLAS computing infrastructure will be centered at the U.S.-ATLAS Regional Center (ARC), at a location yet to be determined. The ARC will be the principal data repository and CPU resource for U.S.-ATLAS. We estimate that at experiment turn-on in 2005, a $5 \times 10^4$ SPECint95 CPU capacity, 50 TB of user disk store, and 250 TB mass store will be required. Planning for this analysis and data center is underway. A working group with members from Labs and Universities has been formed, and this group will participate in the MONARC (Models of Networked Analysis at Regional Centers for LHC Experiment) project, supplementing the current MONARC activity which is focused on simulation and modeling. A long term deployment plan, taking into account the needs of the experiment and the resources of existing facilities, will be developed over the next three years.

With the combination of institutions represented in this proposal, U.S.-ATLAS brings a strong set of capabilities, facilities, expertise, and research alliances to the service of computing for the ATLAS experiment.

Our collaborating institutions are home to some of the premier programs in computing and networking research in the U.S., and have a history of working closely together on these topics. These strong connections to the best in U.S. computer science research will benefit U.S.-ATLAS in many ways. Much of current computing research is devoted to enabling groups of researchers to collaborate across large distances. We will be able to take advantage of the capabilities that these researchers are developing and can even hope to leverage funding from these non-HEP sources by providing test cases and pilot projects from which U.S.-ATLAS and the computing research community can both gain.

The estimated level of computing professional effort over the next three years is as follows. For code development, 7.5 FTE are needed immediately, 10 FTE in FY2000 and 11 FTE in FY2001. For the infrastructure function, 4 FTE's are needed for system operations and user support. In addition, five postdoctoral fellows are needed in FY2000 for reconstruction and simulation code; this should rise to ten in FY2001. The total estimated cost through FY2005 is $40.3M and the operating cost is approximately $7M per year.

The overall project management is not discussed in this document.

# 2 Overview: Project Scope and Goals

This section presents an overview of the scale of ATLAS and U.S.-ATLAS computing. The need for computing infrastructure in the U.S. in the data analysis phase is explained. The scale of the software effort needed to produce a functioning system and the U.S. contribution to it is discussed. The resources needed to support this effort in the immediate term are justified.

## 2.1 The ATLAS Computing Problem

Data taking on ATLAS is scheduled to begin in June 2005 [1]. The event rate is expected to be 100 Hz leading to $10^9$ events per data set which takes one year to accumulate. Each event is approximately 1 MB in size leading to raw data sets of 1 PB; this is reduced by a factor of 10 during reconstruction[2]. Raw data will be reconstructed at CERN. It is therefore convenient to group the reconstructed data into sets of order 100 TB each. We are planning for data sets in 2005 and 2006 to be the same size despite the lower luminosity of the early years because the B physics program requires high statistics and because there will be a desire to obtain data sets with fewer trigger biases than in the high luminosity running.

The ATLAS computing model calls for 500 physicists doing analysis with 150 accessing data at any one time. The U.S. is $\sim$20% of ATLAS, but given the large number of U.S.-ATLAS members who have experience on CDF and D0, experiments with closely related physics goals and analysis strategy, it not unreasonable to expect that U.S. physicists will be in the forefront of analysis and might be a larger fraction of this effort. Therefore we estimate of order 120 analysis users of which 40 are concurrent. Every member of U.S.-ATLAS will need access to the tools; currently there are approximately 250 such collaborators.

The complexity of the ATLAS project and the large data sizes involved require the use of the best possible software design and tools. ATLAS is fully committed to the Object Oriented paradigm. Just as in a large hardware construction project, a large software project requires the right balance of physicists, engineers and other professionals.

ATLAS offline computing is led by the Computing Coordinator. The election of the next Computing Coordinator will take place in June 1999. A recent internal review [3] highlighted the shortage of manpower in the current computing activity. Among the recommendations is that there should exist a Steering Group having as its members persons responsible for core software such as database and control/framework, a person responsible for each of the sub-detector reconstructions, and representatives of ATLAS management. It is also recommended that software deliverables for core software should be negotiated with the Computing Coordinator and that subdetector software such as calorimeter reconstruction code should be negotiated within the hardware subsystem concerned.

The complete or partial reorganization of ATLAS computing that is expected following the internal review is one reason for expanding active work by U.S.-ATLAS computing now. A growing U.S. group can be a positive force as they participate in planning for new directions and overall organization of ATLAS computing.

## 2.2 U.S.-ATLAS Solution

### 2.2.1 Software Development

The present ATLAS software is complicated, and unfamiliar to the majority of collaborators. It is particularly difficult for individual or small groups of would-be developers to contribute, because the startup overhead is so large. A centralized software development effort is less subject to these difficulties, and can also ensure that remote participants can collaborate effectively. Aside from issues of software infrastructure, we see as the function of this centralized group to develop the aspects of ATLAS core software that are U.S. responsibilities (preliminary negotiations with ATLAS management suggest that these will be in areas relating to database, event, and control domains), and to provide development support (largely computer professionals) for physicists developing reconstruction, simulation and other detector specific software. ATLAS estimates that of order 1000 FTE-years will be needed to produce the software [4]. Much of the actual code will be written by physicists most of whom must be trained in the new methodology. Some estimates of the size of individual software projects are given below based on experience from BaBar, CDF, D0 and STAR. Participation by U.S.-ATLAS members in this effort is essential; our collaborators will insist on it. We have an obligation to the future U.S.-ATLAS members who will analyze the data to ensure that the analysis system is well designed and can be used effectively even from a great distance.

In order for the software development to function effectively, some support infrastructure is needed immediately. This includes the hardware needed for code testing, development and prototyping, and the support personnel to integrate and maintain the ATLAS code infrastructure. We propose the minimum level of infrastructure support that we believe is needed now.

Our intent in the development of the software for ATLAS is to gain maximum benefit from the investments in the ongoing efforts at BaBar, RHIC, CDF and D0. We will exploit, wherever possible, the solutions and experience of those groups and actively recruit a number of software professionals involved in those efforts. We expect to collaborate with those groups and with CMS in the development of core software.

It is clear that a large fraction of software development will have to take place in university groups participating in ATLAS. This will be especially true in the case of the reconstruction software related to particular detectors being built by the U.S.-ATLAS groups. However, it is also anticipated that significant contributions will have to be made in development of core software, in which all available expertise will have to be utilized. Coordination of this joint effort is one the important tasks before the management of the US-ATLAS Computing and Software. It is anticipated that funds will be distributed to the universities to support their role in the overall U.S.-ATLAS Computing and Software project.

### 2.2.2 Data Access and Analysis

We will assume a distributed analysis model in which the U.S.-ATLAS [5] members resident in the U.S. rely on a regional center (ATLAS Regional Center –ARC) to provide access to these data and most of the CPU needed to analyze it. The role of regional centers on the LHC experiments is the subject of an ongoing effort, the MONARC[6] project. The vision of a regional center presented here is that of a Tier-1 center in the MONARC notation. Its role is to provide an adequate level of support so that university researchers participating in U.S.-ATLAS will be able to contribute significantly to the software development and eventually to analyze the data efficiently and effectively.

Data will be transported to the U.S. and the U.S. users will primarily access our database and not CERN's via the network. ATLAS intends that all raw data remain at CERN, but certain analyses will surely need some of it. We will therefore have to allow for some raw data resident here in the U.S.. Such raw data could be transported to a (separate) database at the regional center. Access will also be needed to the raw data over the network from CERN. There is an ongoing discussion of a global access to a federated database. The viability of such a database is not known at this time. If implimented, the ARC would need to be a participant. Research is underway (also on CMS) to assess this possibility. The regional center must be able to import, store and make available data sets of 100 TB. This will be stored in an object oriented database, the baseline choice for which is Objectivity. The need to support two releases per year simultaneously, plus the need for some raw data leads to our benchmark of a 250 TB data store at turn-on. It is likely that some of the data will have to be resident on tertiary storage and a hierarchical storage system such as HPSS (High Performance Storage System) will be needed. This system must be fully functional and available to users in 2005.

Typically the analysis takes place intensively on the data from the most recent run. Access to older data will be needed for combining results. A typical user performs an event selection and perhaps reformats the data of interest into a subset that he then works on intensively. These data subsets will vary greatly in size: a few may be as large as several terabytes while others are of order 100 GB. These may be stored as part of the database or kept as files that are not part of the database and are accessed by a separate analysis tool. If the subsets are small, users will want to move them to their home institutions. This transport will take place over the network; the export of data via tape to user institutions in the U.S. is not expected to be needed. These subsets have to be disk resident, implying that 50 TB of user disk space is needed in addition to the common data storage. This too must be available in 2005.

CPU requirements for analysis and simulation are not reliably estimated at this time. The reconstruction system resident at CERN is estimated to require 250 SPECint95-seconds per event (this is about 10 times that for D0 and CDF) leading to a total required CPU capacity of $7 \times 10^4$ SPECint95. DO and CDF expect to have analysis capabilities comparable to those for reconstruction; it is therefore expected that the U.S. will need of order $3 \times 10^4$ SPECint95 (20% of the total) to support the U.S. users. We do not envision any significant reconstruction of raw data being done in the U.S.; to do so would have serious implications for the storage and data transport. There will be heavy simulation requirements as the ATLAS plan envisages all of this simulation being done outside CERN, and the U.S. will have to contribute. The requirement is again difficult to estimate as the new (GEANT-4) based simulation tool is not yet complete and the relative amounts of full and fast (*i.e.* parametrized) simulation are not known. In order to fully simulate a data set corresponding to 10% of a true data set, ATLAS estimates that $5 \times 10^4$ SPECint95 will be needed. Experience from CDF and D0 indicates that simulation and analysis needs are comparable. The U.S. will also have to export simulated data back to CERN for distribution to the rest of the collaboration. We have done this already on a small scale from LBNL where a significant fraction of the full simulation for the ATLAS physics TDR was performed. The detailed CPU requirements will become refined as the ATLAS code is developed and benchmarked over the next few years. As it is vitally important that sufficient CPU be available, for the purposes of this proposal, we will assume a total of $5 \times 10^4$ SPECint95 must be made available in the U.S.

While we believe that it will be more efficient for the large data store to be at one single location, there are several options for the distribution of the analysis CPU. It could be totally centralized so that remote users have very little at their sites. The other extreme is that users make their subsets and then do all their analysis work locally. Large simulation productions will surely need coordination, hence a sizeable fraction of the CPU will have to be under control of a regional center. Personnel costs are by far the largest fraction of the U.S.-ATLAS computing costs, and these are probably minimized with operations and support functions at a single site. Most hardware costs are decreasing so rapidly that it can easily be more efficient to purchase new hardware rather than rely on existing facilities. Nevertheless before beginning the deployment of the first 10% of the production system in 2003, we will make a careful evaluation of the cost effectiveness of a modestly distributed computing environment (LBNL, ANL and BNL and a few large university groups). Progress on distributed computing and remote system management will need to be monitored carefully before the decision is taken. The software development activity in which distributed collaborators are supported from a central site over the next few years will give us valuable experience in this area. Some distribution of resources may also enable leveraging of existing facilities at other university or laboratory sites and more than offset the additional costs.

The issue of platform support is crucial. In this proposal we present a scenario in which ARC runs two platforms: an SMP for analysis and a PC system for simulation. Remote management of external sites will probably require that they run one of these systems. The staff at the center will play a role in testing and supporting of the software for the whole collaboration. Experience from existing experiments suggests the need for one librarian and one full time person per supported platform distributed across the experiment. It may be necessary for the U.S. code developers to have access to all of the ATLAS supported platforms; in this case costs will rise. These issues will be clearly delineated by work that we will undertake over the next year.

Currently ATLAS has a milestone [7] calling for a mock (i.e simulated) data challenge to be completed in Jan 2004. Such mock data challenges have been a vital part of the software development testing and verification on STAR and BaBar. U.S.-ATLAS will need to participate with a significant number of users. Its main function will be to test out the software and refine analysis strategies. We therefore plan on 10% of the production capacity being available in 2003. We will adjust this as the ATLAS milestones become more clearly defined.

To support this analysis CPU and data store, one needs a staff that, in addition to maintaining the system, is responsible for ensuring that the ATLAS code is fully functioning, for exporting the code to individual users, and for aiding those users to operate efficiently. Training in OO methodology and programming tools for physicists participating in the software development must be provided on a short time scale. ATLAS is beginning training in OO design in June 1999. Training in the U.S. can be modeled on it as appropriate. Our software pilot project is an important cornerstone of the training program as physicists will get early hands-on experience with OO coding and tools. It is our intention to hold a meeting in the summer of 1999 at which a GEANT-4 course will be offered and at which we can further assess the needs of the collaboration.

Network/data transport is a key issue. Each data set will need to be updated as it is regenerated from raw data. The ATLAS assumption (conservative) is that the 100 TB set will be reconstructed and sent twice per year. The two methods that data can be transported from Europe to the U.S.-ATLAS Regional Center are: data transport over the trans-Atlantic internet connection or shipping data tapes from CERN to the regional center and shipping the associated meta-data via the internet, or on a tape included in the shipment.

The trans-Atlantic network will certainly be improved, both in bandwidth and in reliability, over the next five years; U.S.-ATLAS must be a strong advocate for the upgrade. However, predictions of how the improvement will occur are fraught with uncertainty. Consequently, the method that assures sufficient bandwidth is the shipping of tapes, and the importing of these tapes into the U.S. ARC mass storage system. This movement will have to be done via air freight if the average network bandwidth utilization of order 100Mbits/sec cannot be sustained. Inside the U.S. however, all data should move over the network. The ATLAS assumption is that each user needs 1 Mbit/sec to his/her desktop if the data reside at the center and the analysis is done locally.

#### 2.2.3 Simulation Activity in the Near Term

It is not clear whether or not significant production CPU and data storage will be necessary for simulation in the near future given that the Physics TDR is almost complete. However, U.S-ATLAS must be able to provide this should it be needed. The site that provides the developer support over the next few years must be able to supply this. We are aware of ongoing radiation background studies using neutron transport codes (M. Shupe and S. Willis) that are limited by the available CPU and could benefit from central support. We expect to respond to the needs of the collaboration in this area.

#### 2.2.4 Code Developer support

The vast majority of simulation and reconstruction software will be written by physicists and not by computing professionals. These physicists will in general be members of the detectors groups responsible for building ATLAS. The model for organizing this software effort, and for insuring that it remains coherent and focused, despite its functional and geographical diversity, is discussed in detail in Section 3.5. We describe here what such reconstruction developers, and the physicists doing analysis, should expect from the U.S.-ATLAS computing project.

The central computing effort must play two separate roles, one related to software development and the other to operations and infrastructure. On the software side, the central group must serve as a repository of knowledge both about ATLAS computing and about OO software technology and practice. It must efficiently transfer this information to the physicists in the collaboration, so that these physicists are free to concentrate on the algorithmic requirements of their code. On the infrastructure side, the central group must provide both facilities support (CPU, access to tapes, disk space, network connectivity) and infrastructure software support (help with installation and code distribution, negotiation and installation of licenses, consulting on core products, etc).

Experience on Babar, CDF and D0 has shown that physicists writing complicated software (in particular relatively inexperienced OO programmers) benefit enormously from design help provided by highly trained computing professionals. The Fermilab Run 2 project, for example, has two FTEs assigned to C++ design consulting. We anticipate the need for a similar function in U.S.-ATLAS. In one model of this activity, groups working on the reconstruction software could request consulting help from the central group. A consultant would then be assigned to the group for some period of time (perhaps a few months) and could provide advice both on general OO design and ATLAS specific implementation issues. Such consultants could also deliver well defined but computationally challenging software components needed by the collaboration.

The infrastructure task is handled by the ATLAS Regional Center (ARC). ARC will provide the ATLAS users with CPU cycles, disk storage, access to the data via robotic tape storage and adequate networking. It is the U.S. repository for ATLAS software, for general HENP software (for example Physics Analysis Tools) and development tools (CASE tools for software design, code monitoring tools to understand code performance, debuggers, etc). The center provides *service* in the form of system managers, user and developer support (including a help desk that can be contacted via phone and email), and centralized code management. Members of ARC will work with ATLAS and CERN to develop and deploy software distribution and validation tools.

The regional center will provide one or more "reference architectures" that are working demonstrations of the interoperability of the sanctioned hardware components, operating systems and software. It will provide advice to U.S. institutions on appropriate choices of hardware. We expect ARC to provide a central clearing house for information, tools and documentation. This group is also responsible for insuring that U.S.-ATLAS physicists have access to training courses and that there is adequate documentation to allow physicists to use the system effectively.

# 3 Software Development Function

## 3.1 Overview of ATLAS Software

ATLAS is committed to the development of object-oriented (OO) software grounded in accepted international standards and practices. The long lifetime of the experiment, the complexity of the required software and the distributed developer environment (we expect $\sim 85\%$ of the software effort to be outside of CERN) all argue for a highly modular system with well defined interfaces.

ATLAS has addressed the issue of software design in its Computing Technical Proposal and had codified its methodology in a well defined “ATLAS Software Process” (ASP). This process includes formal design reviews and required stages of documentation. Details of the ASP are currently under review, with the aim of streamlining the process; we expect, however, a continued emphasis on modularity, maintainability and documentation.

To encourage such modularity and to allow the division of software tasks among the collaborating institutions, ATLAS has partitioned the software components into semi-autonomous “domains.” Work within each domain is coordinated by a domain architect, with well-specified and controlled interfaces between domains adjudicated by a Domain Interaction Group. Examples of domains in the current ATLAS software include: Control, Reconstruction, Graphics and Simulation.

A significant fraction of the OO software effort on ATLAS has been devoted to the evaluation of external components of the software environment. Much of this work has been coordinated through CERN-supported R&D projects. The following baseline design choices have been made by ATLAS:

- **Object Database Management System (ODBMS)**
  ATLAS expects to store all data (raw, reconstructed, analysis, calibration) using an object database. Advantages of this approach include the ability to manage data from a logical perspective using OO techniques and the ability to store an event hierarchically with some frequently accessed parts of the event disk resident while less often used information is stored on tape. Although the final choice of ODBMS vendor will not be made until 2001, Objectivity is currently the baseline. It is ATLAS policy that database access occur using a risk-averse strategy that shields the application programmer from the choice of vendor. This policy can be implemented, for example, by providing an adaptor layer that maps between the transient view of the event seen by the application and the persistent view that is stored in the database.

- **Mass Storage System (MSS)**
  The vast majority of the ATLAS data will be stored on tape and will not be disk resident. The hierarchical mass storage system is responsible for migration of data files between tape and disk. Because the MSS must be interfaced to the ODBMS, the characteristics of this system are relevant to the core software developers. CERN currently is using the High Performance Storage System (HPSS). This system has been interfaced to Objectivity.

- **Core Development Language(s)**
  At present OO reconstruction and simulation software is being written in C++. This choice is consistent with industry and HENP practice. There is a significant development effort using Java in graphics and event-display applications.

- **Simulation**
  ATLAS will use the Geant-4 toolkit for event simulation. Work has begun on a preliminary description of the detector using Geant-4 and several ATLAS-specific classes on the use of Geant-4 have been held. Significant effort will be required to validate the Geant-4 simulation against ATLAS testbeam data.

- **Code and Configuration Management**
  A code and configuration management system is currently operational on ATLAS. Like most current experiments, ATLAS has chosen CVS as its code repository. An ATLAS-developed variant of the SoftRelTools (SRT) package originally developed for Babar is in use for library building and configuration management. Frozen releases are built and are distributed to remote sites via AFS.

- **Platforms and Operating Systems**
  ATLAS software is supported on a number of Unix platforms. HP-UX is the current platform used by most of the CERN-based developers. ATLAS expects Linux and Sun-Solaris to become the main Unix platforms over the next few years. Re-evaluation of the list of supported platforms will be necessary at regular intervals.

Much of the ATLAS software currently in the CVS repository is legacy Fortran developed during the preparation of the ATLAS proposal and extensively exercised and improved during the preparation of the detector TDRs. We anticipate development of this legacy code to cease with the submission of the Physics TDR in May 1999. It is expected that much of the manpower currently devoted to the maintenance of this code will move on to development of the OO software system.

## 3.2 Pilot Project

A first step in US-ATLAS computing will be a "Pilot Project" aimed at provided a monitoring and analysis system for test-beam data to be taken for the Tilecal detector during the summer of 1999. The initial goal will be to place testbeam data into an object data store both at CERN and LBL, and to allow physicists to carry out calibration and analysis of this data using C++ code and OO-based software tools. The Tilecal group has been taking test-beam data for each of the last several summers at CERN, and there is considerable experience in data-taking and Fortran-based analysis. Consequently the goals of an analysis system are well-defined, and we will have group of highly motivated physicists to use and evaluate the result. The goals of this project are at least three-fold:

First, we believe this will be an important learning experience for the software developers. By actually developing and commissioning a full albeit small analysis system that will be used by physicists, they will be forced at a very early stage to confront problems at a number of different levels: data structures, data stores, access to data, incorporation of calibration data, interfaces to GEANT-4 and to analysis packages, coding standards, and documentation. This will be enormously useful in bringing attention to problem areas, and in introducing everyone involved to all aspects of the system. It will provide an essential "bottoms-up" component to software planning. Secondly, this will give ATLAS physicists an immediate introduction to C++, object-oriented design, objectivity, and new analysis tools. It will provide other test-beam groups with a good model, which will be quite important in shifting the collaboration from Fortran code to the object-oriented framework that will be used in ATLAS. We believe the Pilot Project will be important in bringing physicists, as well as software developers, into the development of an ATLAS analysis system. Thirdly, this system will be an important tool in the analysis of Tilecal test-beam data – this essential goal is in fact what drives the Milestones listed below. We have discussed this project in detail with Marzio Nessi (Project Leader of the Tilecal group) and the Tilecal group, and believe that we can interface well with testbeam plans and activities. The Tilecal group will continue with test-beam studies in the summers of 2000 and 2001, and we expect this system to be used, and to continue to evolve, during the next three years.

The Pilot Project will be especially important in allowing new analysis systems such as LHC++ and ROOT to be applied to data analysis, both for the analysis capabilities they provide and to allow a wider evaluation of these packages by users. Current participants in the pilot project are from ANL, U. Chicago, Duke, and LBNL.

## 3.3 Core Software

In ATLAS, as in most HEP experiments, software is divided into core software and non-core (detector specific or physics analysis) software. U.S.-ATLAS expects to contribute to both these areas and has specific experience and expertise to play a leadership role in several core software domains.

Core software is typically written by computing professionals rather than physicists and does not pertain specifically to a particular detector or to a physics program. It provides the operating environment for the software modules written by physicists, and supplies mechanisms for control of those modules, for data input/output to them, and for communication and coordination between modules written by different people and groups. Core software also provides other functionality that facilitates reconstruction and physics analysis tasks and that insulates the user (i.e. the physicist) from the underlying details of the operating system, data format, communication protocols, *etc.*

Software development in the core areas is both crucial to the work of ATLAS physicists and presently drastically under-staffed. Because it provides the environment into which a physicist must integrate her own code, it must be developed and made operational early in the software process. The design and development effort for core software must, therefore, be heavily front-loaded. It is not unreasonable to expect that the next several years will require more work on core software than on detector-specific and physics software.

Because core software is both a particular strength and a vital interest to U.S.-ATLAS physicists, we propose to take lead responsibility for at least one core software domain and to contribute significantly to a second core domain.

The selection of U.S.-ATLAS computing's core domains of responsibility will be made in close consultation with the overall ATLAS Computing Coordinator and should result in an explicit agreement on the U.S. role and deliverables. We identify below two core domains where there is considerable U.S. strength and interest.

### 3.3.1 Control/Framework

The earliest control framework for ATLAS (ARVE - ATLAS Reconstruction and Visualization Environment)[8] was written by a U.S. collaborator, Toby Burnett of University of Washington. ARVE is essentially an event pump and a way of chaining analysis modules for serial event processing. It provides a graphical user interface and elementary visualization capability as well as some basic geometry classes. To address some of the shortcomings of the original ARVE implementation, a component object network model has been proposed. This model has been described in detail in Lassi Tuura's master's thesis [9]. No functional implemention of this model has yet been used with genuine physics code. Members of the U.S. ATLAS computing project have volunteered to evaluate the current ATLAS control domain software by attempting to use it with LAr reconstruction code. At present, no active software development effort on component networks or ARVE is being conducted.

We believe that an evaluation of the current ATLAS control and framework structure, along with a "market survey" of comparable systems from other experiments is a necessary first step in the development of a long-term strategy for the ATLAS control framework.

Some of the primary tasks are to control efficiently the flow of the event reconstruction, and the I/O of event, calibration and control data, efficiently using the CPU and I/O resources available for the production batch jobs. To process each event, analysis and service modules (including I/O modules) will be connected in multiple execution paths, corresponding to different event classifications defined by the experiment trigger. Each path will have independent filters and I/O modules. The framework will make sure that modules common to more than one path will be executed only once, and will control the transitions of the modules through a finite, configurable set of states (start-up, run begin, event begin, etc.).

Initially, the framework will be used by the ATLAS software team as a development tool. As ATLAS gets closer to data-taking the number of "end-users", many of whom will actually contribute substantial amounts of code to the plug-in modules, will ramp-up to several hundred.

The framework must provide the user with an interactive shell offering the full functionality of the production jobs, including the ability to define execution paths and configure modules, the dynamic loading of modules, and the ability to stop event analysis in the middle of an execution path on the basis of the information provided by an event selection or filter module. It must offer a consistent interface to pass module output data to the event display and to any of the supported physics analysis packages.

We believe that this framework is essential to providing the physicist with the ability to analyze and visualize the data in the most productive manner possible. We envision the use of a format-independent data conduit that will permit the use of several different analysis/visualization packages, and that can be expanded in scope as new products become available. It will offer the ability to deal with complex objects as well as simple literals, and interface directly with the user's analysis code. It is not intended to be a replacement for the likes of PAW and ROOT, but rather an interface layer between analysis code and a visualization package, permitting the user to chose whichever final format is desired. By calling on routines provided by this package, the user will be able to histogram simple variables, object members, and even entire objects, in multiple dimensions, with minimal additional user burden. Upon completion, the histogrammed data will be saved in whichever format was selected by the user, including customizable formats, allowing continued analysis or visualization at the user's discretion.

The use of such a data conduit will free physicists to pursue analyses in a less rigid fashion than is currently possible. By facilitating such freedom, we can start to make industry standard tools more accessible to the average physicist, as well as allow specialized tools to be developed for particular applications, supplementing general purpose tools such as PAW and ROOT with highly focused ones. This will allow a new freedom of choice and add flexibility that has, until now, been lacking from the HENP domain.

### 3.3.2 Database

ATLAS (like CMS) has chosen to store its data in an Object Database. If sucessfully implemented, this choice provides significant advantages for the physicists analysing data. It removes the need to buffer full events into memory and reads only those parts of the event the user has requested. This significantly improves the I/O rate for objects of interest. It presents a full Object-Oriented description of the data, including the ability to store pointers and associations. It allows databases that are widely distributed at different sites to be viewed as a single logical unit, the "federated database." And finally, when used in conjunction with a tape store, it allows for transparent access to all event data.

We propose U.S. contributions to ATLAS database efforts in three areas:

- **Infrastructure for distributed database development, and for wide-area access to physics data**—without this infrastructure, serious contributions from outside CERN are difficult, if not impossible;
- **Design and implementation of the control/database interface, and of generic database components needed to support that interface**—these components are necessary for the deployment of the control framework whose development the U.S. proposes to lead;
- **Database scalability**, which is an area of particular U.S. strength, and one that must be addressed in the design stages, even though, insidiously, design failures here may not become apparent until significant amounts of data arrive.

**Database infrastructure** will be our earliest focus—without much of this work, further development cannot proceed. The U.S. tasks could include

- Establishment of mechanisms and precedents by which large databases may be delivered to outside institutes;
- Development of database coding and interaction rules and standards akin to C++ coding rules;
- Devising of a risk-averse strategy for use of the Objectivity/DB database product within ATLAS, and implement it in database coding standards;
- Definition of a reference database to contain standard ATLAS schema for use by developers, based on the experience of BaBar and other U.S. experiments;
- Provision of a mechanism by which a reference database may be built remotely, equivalent to building a software package checked out of a repository;
- Establishment of protocols for sharing schema, for avoiding schema conflicts among developers, and for importing schema from outside.

Because it is fundamental to subsequent development, most of the milestones related to this effort appear early in the our software plans. See Table 3.

**Control/database interface** milestones will be timed to support and enable delivery of the proposed control framework, whose milestones are summarized in Table 2. Early tasks include

- Design and implementation of an approach to maintenance of database and transaction context as control flows from component to component;
- Design and implementation of the means necessary to refer to, locate, and move persistent data into and out of software components that use the control framework;
- Participation in the definition of the approach to transient/persistent mapping as part of the global architecture effort, and provision of the control mechanisms needed to support this approach.

This is a highly collaborative area for development, and will rely on interactions, not only with control framework developers, but with those charged with defining the shapes and (persistent) states of physics objects—not just events and their properties, like tracks and hits, but also calibrations and geometries—and their associations to other objects.

*Database scalability*, and, more generally, architectural scalability, are properties that cannot be added like one more new feature to software after it has been engineered—scalability must pervade system design from its inception. Much of the proposed effort, therefore, involves collaboration with other efforts for which we are not proposing that the US take primary responsibility, though there are certainly many specific components of approaches to database scalability that we do propose to deliver, as well as component scalability assessment efforts that can be undertaken independently.

It should be understood, then, that a portion of this effort is an ongoing *consultancy* role—we may not take primary responsibility for event collection management, for example, but we *will* participate in ensuring that the design and implementation are scalable to petabytes of data and billions of events. Design scalability efforts must begin early, and persist through the lifetime of the project.

Evaluation and assessment of the scalability of specific technologies and approaches will be accomplished in the following ways:

- By understanding the experience of current and near-term experiments like BaBar, RHIC, and Fermilab Run II;
- By means of common projects with experiments that plan to use the same technologies (*e.g.*, jointly with other LHC experiments via RD45 for Objectivity/DB—efforts in which the proposers already have a track record of significant involvement);
- By means of the calorimeter testbeam pilot project, which will, for example, provide information about the performance viability of maintaining distinct transient and persistent data models;
- Via other short- or medium-duration projects that, like the pilot project, will be designed to evaluate approaches to data handling in the context, not of artificial tests, but of applications that address genuine physics needs.

These scalability assessments must be accomplished on timescales that match ATLAS decision timetables, *e.g.*, for database selection.

Delivery of specific components for scalable data handling (e.g., prefetching mechanisms for data on tertiary storage and parallel query and analysis capabilities) will be essential to the success of ATLAS computing, but not to the first years of software development. These efforts account largely for the increase in database effort needed in the years approaching 2005. See Table 3.

### 3.4 MONARC

MONARC is a CERN R&D project created jointly by the LHC experiments. Its goal is to develop the simulation tools and guidelines to allow the LHC experiments to model their future computing systems based on a distributed data architecture. Krzysztof Sliwa (ATLAS) and Harvey Newman (CMS) led the effort of preparing the proposals. The project was approved by LCB in December of 1998. The primary motivation is to explore ways to maximize the contributions to physics analyses by physicists distributed geographically all over the world. Fundamental questions which need to be answered are

- How many centers of what size and equipped with what resources will optimize the analyses throughput while minimizing the overall cost ?
- How the available network bandwidth constrain influence the distributed data architecture models.

A realistic simulation, validated as much as possible with real-time measurements on existing systems and networks is needed.

The Simulation and Modeling Group is led by Krzysztof Sliwa. Included is work on the optimization of system architectures. A major milestone was the development and construction of the first round models. A toy model of a CS2 computer farm, a system which has been modeled in the past and whose performance parameters were known, has been developed to demonstrate the ease and flexibility with which Java-based tools can be applied to the problem. A simple model of a disk-server with multiple processed writing into an Objectivity/DB database has been created to allow comparisons with real systems in order to verify the logical model in the simulation.

A model of a single center, with realistic description of interactions between the independent processes, and with the realistic data model built-in into the model of the database structure, exists now. A program of test runs with varied loads and access patterns is currently in preparation with the primary goal of defining a more complete set of output parameters with which to characterize the performance of the system. At the same time further development work on the simulation tools is under way. Discussions with and members of the the RD45 project at CERN, have resulted in a more refined logical model of AMS (Objectivity/DB) servers and their interactions with other passive and active objects being created.

With those elements in place, one can connect a number of centers, creating a first version of the MONARC model of the entire distributed computing system. We expect to have this first model in place by the end of April.

A series of measurements will begin soon to validate some elements of the model itself, as well as to provide a number of parameters which are input to the model. The plan is to create a federated Objectivity/DB database with data replicated in Autonomous Partitions at several, geographically remote, sites, connected via networks with wide range of available bandwidth. LBL could contribute important measurements in this area, especially those which would address questions related to the interaction between the hierarchical storage manager (HPSS) and the Objectivity database.

This effort is currently funded for 1.0 FTE of software professional. We anticipate that this level will be maintained.

### 3.5 U.S.-ATLAS Reconstruction and Simulation Software

A major part of the ATLAS computing effort will be the writing of the event reconstruction code for each of the ATLAS subdetectors. This will differ from core software projects in a number of important respects, and will require a somewhat different organizational structure. The reconstruction software effort will be dominated by physicists from the subdetector groups rather than by software professionals. It will involve a much larger complement of university physicists, including postdocs and students, and consequently will be more geographically dispersed than the core projects. Its scheduling will have to take account of the manpower limitations and hardware responsibilities within each of the subdetector groups.

Although the ATLAS reconstruction effort will necessarily have to be organized in part through the subdetector groups, it is essential that there be a strong overall organization to foster close communication among the different subdetectors, since decisions regarding data and code structures in one area will impact the structure of the entire system. In addition, there must be a central software design group to develop the overall structure of the analysis system. Because of the dispersed effort, it will be especially important to provide strong central support, well-designed training programs, good documentation with examples, effective communication channels, and a coherent overall structure.

The scope of the reconstruction effort must also include Monte-Carlo simulations, that will be essential, not only in data analysis, but in the development of algorithms for event reconstruction even in the early stages of reconstruction work. Since simulation development must extend down to the subdetector level, it is natural to include the full simulation effort within the reconstruction organization.

To oversee this work, it is expected that the overall ATLAS collaboration will set up a Reconstruction and Simulation Group which will be responsible for the central design of the analysis system and which will oversee the development within the different subdetector systems. To this end, each ATLAS subdetector is expected to appoint by mid-May an offline coordinator, a reconstruction task leader, a simulation task leader, and a database task leader.

Similarly U.S.-ATLAS will have to provide a structure that is responsible for assessing the needs of the developers and for providing coordination/allocation of software professional resources. Although the organizational details may change as the ATLAS computing organization becomes more fully developed, the responsibilities of this structure will likely include:

- Facilitating participation of all U.S.-ATLAS institutions, especially university groups, in the development of reconstruction, analysis, and simulation programs;
- Advising the U.S.-ATLAS Associate Project Manager for Computing about the allocation of funding, including support for manpower and hardware, to ensure that adequate resources are available to meet milestones;
- Assigning a core group of CS specialists who will actively participate in overall program design, in approving software developments of subdetector systems, and in providing guidance and expertise as needed.
- Providing input on the training requirements so that the training described in section 4.3 is attuned to the needs of physicists.

Because of strong U.S. involvement in many other large HEP experiments, and because of the experience and computing strengths of U.S.-ATLAS, we expect that the U.S. groups will play an especially important role in the subdetector reconstruction and analysis efforts, a role which can be significantly enhanced with a well-organized U.S. ATLAS computing organization.

This reconstruction and simulation effort will be a massive one. Its work will include the full calibration and alignment procedures for all subdetectors; the development of algorithms for extracting energy, momentum, and particle ID; the reconstruction of jets and the determination of missing energy; the detailed investigation of systematic biases in trigger selection; extensive studies of rate-dependent effects in all subdetectors; and, in general, the full investigation of the physics reach of ATLAS.

Because this work will actively involve all university groups, and because of the importance of the extensive educational activities that it encompasses, this area may be an appropriate focus for support by the National Science Foundation. The NSF has highlighted both information technology and education (especially education that will serve to broaden the skills of physics students) as areas of special emphasis over the next years, and these goals coincide very well with the focus of the proposed reconstruction and simulation activities. We expect that much of the reconstruction and simulation work will be carried out by postdocs at universities. NSF support of these postdocs would seem to be a natural way of moving in just the directions that the Foundation has chosen to emphasize. Although the U.S. hardware responsibilities will continue for some years, there is urgent need now to train additional U.S. postdocs in OO software techniques so that they may contribute soon to the reconstruction and simulation work.

It is also important to note that there should be no artificial distinctions in support of this reconstruction and simulation area between groups whose research support comes from NSF and those whose support comes from DOE. Support for this important effort should be able to accommodate all universities and all national lab groups that desire and are able to participate.

As we expect that most of the software in this area will be written by physicists, it is prudent to plan for a number of physics postdocs in this area. In FY2000, approximately one additional postdoc per subsystem, *i.e.* five for U.S.-ATLAS is the minimum number that is will be needed. This number should be expected to double in FY2001.

## 3.6 Software Development Milestones and Resource Estimates

Summaries of the milestones for software development are shown in Table 1 through Table 3. An estimate of the necessary FTEs (this number includes only computing professionals, not physicists) is given in Table 4.

| Date | Milestone |
|---|---|
| 17 Mar 1999 | Project begins (DONE) |
| 12 Apr 1999 | Persistent Model selected (DONE) |
| 15 Jun 1999 | Phase one complete: new transient model implemented |
| | Simple analysis tools connected ready for beam |
| Aug 1999 | Connection to GEANT4 implemented |
| Sep 1999 | New persistent model implemented |
| Oct 1999 | Expanded Suite of Analysis Tools available |
| Dec 1999 | Combined testbeam data models complete |
| Mar 2000 | Combined testbeam analysis components available |
| Apr 2000 | Project ends; tools move to control and database; |
| | calorimeter specific code moves to reconstruction |

Table 1: Milestones for TileCal pilot project

| Date | Milestone |
|---|---|
| Dec 1999 | Requirements Document Complete |
| | Basic architecture specified |
| Feb 2000 | Interface definition language and framework language chosen |
| | Scripting language specified |
| Oct 2000 | First package released to application developers (Alpha) |
| Apr 2001 | Evaluation of Alpha release complete and documented |
| | Freeze module interface |
| | Freeze core architecture |
| Oct 2001 | Second version released to ATLAS users (Beta) |
| Apr 2001 | Evaluation of Beta release complete and documented |
| | Freeze architecture for distributed applications |
| Oct 2002 | First production release with full functionality |

Table 2: Milestones for Analysis Framework (Control Domain) project.

| Date | Milestone |
|---|---|
| Jan 2000 | Infrastructure for distributed database developers deployed |
| Oct 2000 | Infrastructure for distributed access to physics data deployed |
| Oct 2000 | Release of database infrastructure in support of control framework's alpha release |
| Jan 2001 | Scalability assessments of candidate data-handling technologies completed |
| Oct 2001 | Release of database infrastructure in support of control framework's beta release |
| Oct 2002 | Release of database infrastructure in support of control framework's production release |

Table 3: Initial milestones for the Database project.

| | FY99 | FY00 | FY01 | FY02 | FY03 | FY04 | FY05 | Total |
|---|---|---|---|---|---|---|---|---|
| Pilot | 3 | 2 | 0 | 0 | 0 | 0 | 0 | 5 |
| Control | 2.5 | 4 | 4 | 4 | 3 | 3 | 2 | 22.5 |
| Database | 1 | 2 | 4 | 4 | 6 | 6 | 6 | 29.0 |
| Collaboration Design Support | 0 | 1 | 2 | 2 | 3 | 3 | 3 | 14 |
| MONARC | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 2 |
| Other Core | 0 | 0 | 1 | 2 | 2 | 2 | 2 | 9 |
| Total | 7.5 | 10 | 11 | 12 | 14 | 14 | 13 | 81.5 |

Table 4: Summary of Software Professional FTE effort. The sum through 2005 is also shown.

### 3.6.1 Tilecal Pilot Project

Work on the pilot project has begun and is planned to last for another year. The project is divided into three phases:

- **Phase 1: Early Tilecal**
  The goal of this phase is to provide an online monitoring framework for the 1999 testbeam. This phase will recycle existing software as much as possible. The major deliverable for the U.S. is the design and implementation of a transient data model for the TileCal data. This transient model will be connected (via an adaptor class) to the existing Objectivity database schema developed at CERN.

- **Phase 2: Extending the Structure**
  The goal of this phase is to redesign the Objectivity database schema to better match the long term ATLAS design. In particular, the redesigned schema will allow for delivery of small subsets of the data (selection by run and by parts of the event). An interface to GEANT4 will be provided in this phase and an evaluation of available Physics Analysis Tools will be underway. A mechanism for storing calibration data will be developed.

- **Phase 3: Multidetector Testbeam**
  In this phase the persistent and transient data models will be extended to include the Liquid-Argon calorimeter data and calibrations. Refinement of the data model to reflect development of the ATLAS event model will also take place at this time.

During the period described above, we estimate the need for 3 FTEs of computing professional support. The tools developed in this project, as well as the manpower, will move to the Control and Database projects. The TileCal specific software will become part of the Reconstruction and Simulation project.

### 3.6.2 Control/Framework Project

Milestones for the development phase of the Analysis Framework project are shown in Table 2. After an initial period of gathering requirements, a development cycle of $\sim$ 3 years will be necessary to complete a baseline product.

- **First Release (Alpha):**
  This release will demonstrate the basic functionality of the package. We expect it to be adequate for use by application code developers, but not for casual users. A preliminary version of the application interface will exist, but these interfaces will not yet be frozen. The functionality in this release will include: dynamic loading of I/O and analysis modules, run-time specification (via configuration script) of the order of module execution, and the ability to process events through multiple analysis paths and to discontinue event processing based on the filter decisions signaled by the modules. The release will include an interface to at least one Analysis Tool Package that allows for the creation of histogram/ntuple files but it may not support interactive use of that Analysis Tool from within the Framework. Framework and module configuration information will NOT be permanently recorded to a file or database in this release.

- **Second Release (Beta):**
  This release should be adequate for early testing by physicists who are not core members of the application development team (in other words users who wish to analyze reconstructed data rather than developers of reconstruction code) in addition to the application developers. We expect all I/O interfaces to be finalized by this time. The additional functionality relative to the Alpha release includes; support for recording Framework and Module configuration information to a flat file or database and the ability to reconfigure from that file; an interactive interface to at least one Physics Analysis Tool; support for an abstract interface to additional graphical applications (such as event displays); and support for a set of module and Framework and monitoring tools that allow developers to gather statistics on CPU usage, memory usage, *etc.*. This release need not support multithreading nor support tools for distributed configuration across multiple clients. A Graphical User Interface (GUI) is not a milestone in the Beta release.

- **First Production Release:**
  This release is expected to be fully functional and adequate for production operations. It will support multithreading and will include a Production Manager that can be used to control distributed applications. GUI support will be included.

We have scheduled this work to begin immediately, since the Analysis Framework is already a critical path development item for ATLAS. During the development cycle, we estimate a need for 4 FTEs of software professional support. Once the baseline product is released, the support load will decrease. We have, however, included significant maintenance support in the period from FY2003 through FY2005, since this period will be at the peak of ATLAS software development, and the potential exists for significant evolution of the ATLAS software environment.

### 3.6.3 Database Project

The time evolution of the database activities as having three phases:

- **Phase One** (years one and two) of the database effort addresses establishment of the infrastructure for wide-area database development and use. This phase also encompasses the design of components of the control/database interface, and is characterized as well by participation in design activities of other software domains to ensure the scalability of their components to massive amounts of data. Participation in collaborative efforts with other experiments to assess scalability of specific technologies, and involvement with the calorimeter pilot project to evaluate the scalability of specific approaches, also occur in this phase.

- **Phase Two** (years three and four) shows an increase in database effort, largely to implement and deliver the database components needed to support the control framework. Connections to the database will also need to be provided for other software domains and detector-specific codes that are, in this timeframe, connecting to the control framework.

- **Phase Three** (years five and six) is when U.S. database efforts peak. In this phase, we must support reliable storage and delivery of significant amounts of data to physics codes. Because petabytes of data begin to arrive at the end of this phase, components that directly address scalability—prefetching from tertiary storage, cache management strategies for concurrent queries, support for parallel database population by reconstruction codes and parallel queries by analysis codes, to name a few—must be implemented in this phase.

### 3.6.4 Additional Software Activities

In addition to the projects described above, we have included in our resource estimates several other tasks. The MONARC project is expected to continue through FY2000. One FTE is assigned to this project. As U.S. physicists on ATLAS begin work on reconstruction and simulation code, there will be a need to provide software professional support for their designs efforts. The model for deploying this support is described in Section 3.5. We anticipate 1 FTE of such support in FY2001, with a ramp-up to 3 FTEs by FY2003. We expect to contribute to the overall design of the ATLAS software architecture by participating in software design reviews, code reviews and tool evaluations and by serving on various *ad. hoc.* committees within the collaboration. (For example, the recently formed 9 member architecture design committee will include 2 members from LBNL.) We have estimated the resources required for the sum of those activities not separately budgeted as line items to be roughly 2 FTEs beginning in FY2003. This is shown as "Other Core" in the Table 4.

# 4 Infrastructure

We use the term infrastructure to denote the computing hardware, software, and technical personnel that do not contribute directly to ATLAS software, but make the development, distribution, maintenance, and use of that software possible. U.S.-ATLAS Computing will have several distinct phases over the next decade and beyond. These are:

- **Development Phase** (Now-June 2002). Software development and planning for deployment of computing resources.
- **Partial Deployment Phase** (June 2002-January 2004). Continue software development and complete partial deployment (10% of final system) of computing resources for the Mock Data Challenge.
- **Full Deployment Phase** (January 2004-June 2005). Complete software development and deployment of production computing resources.
- **Operations Phase** (2005+). Computing resources are utilized for data analysis, with software development continuing at a reduced pace.

These dates will be revised as necessary as the ATLAS computing schedule evolves.

In our planning for infrastructure, we have made the following assumptions:

- The needs of the collaboration for CPU, data storage, and networking in the development phase (about the next three years) can be met largely with existing facilities at ANL, BNL, LBNL (NERSC), and at the universities with modest additional hardware.
- However, incremental support personnel to facilitate software development and maintenance by the broadly distributed engineers and physicists creating and using code are essential during the development phase, starting immediately. To optimize cost-effectiveness, these personnel should be located predominantly at one site during this phase.
- The final location(s) of the hardware and support/operations personnel is an open issue and need not be fixed for about two years, in preparation for partial deployment of the Regional Computing Center. The hardware and personnel may eventually be located at a single site or multiple sites to maximize the utilization of resources available to the Collaboration. For simplicity, however, we later present a cost estimate based on location at a single site.

We summarize in Table 5 the projected CPU power, disk storage, mass storage and support/operations personnel for each of the four phases defined above.

## 4.1 CPU and Storage

HENP experiments currently engaged in data analysis and others contemplating imminent data acquisition each have their own computing architecture and/or set of supported architectures. Such architectures typically include clusters of workstations or SMPs.

| | Development Now-6/2002 | Partial Deployment 6/2002-6/2004 | Full Deployment 6/2004-6/2005 | Operations 6/2005 |
|---|---|---|---|---|
| CPU (SpecInt95) | 100 | 5000 | 50000 | 50000 |
| Disk (GB) | 250 | 5000 | 50000 | 50000 |
| Mass Storage (TB) | 0 | 50 | 250 | 450 |
| Management Personnel (FTE) | 1 | 1 | 1 | 1 |
| Support Personnel (FTE) | 2-3 | 4 | 5 | 5 |
| Operations Personnel (FTE) | 1-2.5 | 6 | 8.5 | 8.5 |

Table 5: Summary of hardware and personnel requirements for the U.S.-ATLAS Regional Center with the assumptions given in the text

The capability of computer hardware is changing at a tremendous pace. It is difficult to predict with precision and certainty now what will be available a few years in advance of our production needs. We present here one possible scenario for hardware acquisition for the U.S.-ATLAS Regional Center with the understanding that our plans must be continually reviewed and revised as new hardware becomes available and new technologies emerge.

### 4.1.1 Development Phase

The "Development Phase" will involve the rapid coding, testing, debugging, and deployment cycle characteristic of a software development project. The hardware requirements during this phase are different from those of a production system in that the amount of interactive work (editing, compiling, running debuggers, etc.) is proportionately much higher than the level of batch processing.

There must exist systems large enough for meaningful tests, but since hardware purchased and available during this phase will be obsolete by the time true production begins, it would be wasteful to purchase more hardware than immediately necessary for programmers compiling and testing code and for physicists running detector simulations.

We propose to utilize the existing resources of the Collaboration, with some modest enhancements to the PDSF. The current PDSF architecture is listed in detail in Appendix A and includes the following components;

- Sun Solaris special purpose servers
- Sun Solaris interactive development workstations
- Intel Linux batch analysis workstations
- Intel Linux interactive development workstations
- Linux NFS data vaults
- Solaris Veritas NFS file volumes

PDSF is a shared system operated by NERSC for the High Energy and Nuclear Physics experimental community. Its other HENP users include STAR and CDF collaborators. Of its total current capacity, the ATLAS collaboration can expect to have access to:

- CPU = 159 SPECint95
- Local Disk Space = 120 GB (average)
- Data Vault Disk Space = 110 GB

We expect that the CPU power necessary to do simulations, software development, and meaningful tests will not exceed, nor even approach that currently available during the development phase. Hence we project no purchase of CPU hardware in this phase, except for replacements of failed equipment and the purchase of ATLAS reference systems for developing and testing ATLAS software as well as testbed systems for eventual production hardware.

Testbed systems are production-like systems that precede the real production systems and allow support and operations staff to become familiar with the hardware. Prior experience with these technologies is essential to put them into production in a stable configuration.

The disk space required to analyze calorimeter testbeam data will be on the order of twice the raw data volume. Assuming a data volume of 100 GB in 1999, this implies 200 GB of disk space for data, plus space for code development, ATLAS-specific software tools, the U.S.-ATLAS computing code distribution set, etc. We propose purchasing 200 GB of space for test beam data and using the existing PDSF data vault space for non-data requirements.

We expect little actual use of HPSS during the Development Phase except for software development and testing as we expect that test beam data will reside primarily on disk. HPSS was developed to address in particular the data storage problems and data access patterns of users of mass storage systems at supercomputing centers, and for the purpose for which it was designed HPSS has worked well. It has been a concern that the HPSS developers are not sufficiently aware of and concerned with the special data access issues of experimental HENP. We believe this can only be remedied by an HPSS developer at NERSC charged specifically with addressing HENP data access issues, and responsible for bringing those issues into the mainline HPSS development process. However, because of the time scale of the HPSS development cycle, this HENP-oriented HPSS developer must be hired and must begin working during the development phase for U.S- ATLAS computing.

#### 4.1.2 Partial Deployment Phase

In this phase we would begin to acquire the hardware needed to support full-scale analysis and simulations. We plan to have 10% of the full system capacity deployed in time for completion of the Mock Data Challenge by January 2004. By late 2003 we would have in place 5000 SpecInt95, 5000 GB of user disk, and an HPSS dedicated capability.

#### 4.1.3 Full Deployment and Operations Phases

As stated in Section 2.2.2, we expect to supply $5 \times 10^4$ SPECint95, 50 TB of user disk space and 200 TB per year of hierarchical storage and to support 40 concurrent users during the first year of the operation phase. The configuration of the system at that time cannot be predicted with any degree of certainty. In estimating the cost of CPU and disk storage we have extrapolated from current costs using Moore's Law [10] and assumed that the cost of CPU and disk falls by a factor of 1.5 per year. The CPU cost is kept constant after 2004 where the feature limit of current technology might be reached [11]. These numbers are somewhat more conservative than the ones derived by Richard Mount [12] which are often used in extrapolations. As a baseline we assume that 50% of the CPU is provided by commodity hardware and 50% by more expensive SMP's. This ratio may change, but it gives us a reasonable capability for I/O without sacrificing the cost savings advantage of commodity hardware. It also corresponds to current usage on CDF and D0 where commodity farms are used for simulation and SMP's for data analysis. For comparison, an enterprise-class Sun SMP machine is currently about eight times more expensive per SPECint95 than a commodity machine (Linux or NT box).

We have assumed that the hierarchical storage is HPSS, in accordance with the CERN choice for ATLAS, and have extrapolated from the current costs. The largest single cost is the robot. We have assumed that we will use STK robots and have taken their current costs as a basis for extrapolation to later years. Similarly we have also costed the tape drives and individual tapes at their current prices. To extrapolate to later years, we have then assumed that the media capacity and the tape drive I/O bandwidth will increase by a factor of approximately eight between now and 2005. We believe these assumptions to be conservative. Although tertiary media costs have in the past fallen more slowly than CPU and disk costs, and can be expected to continue to fall more slowly since they are not subject to the same level of commodity market pressures, technology advances will likely result in some significant reduction in price over this time period, as we have assumed in these estimates.

We will be required to provide these hardware resources on a 24 hour per day-7 days per week basis and have estimated the manpower needed to sustain the CPU and data store accordingly. If we are operating a stand-alone CPU system, we estimate that 6 FTE's will be needed to provide this coverage. In addition we have budgeted 5 FTE's doing tasks associated with the ATLAS code and with ATLAS users (see Section 4.2.3).

If the HPSS site is LBNL(NERSC), we expect to be able to manage the storage with 2.5 FTE's which would be incremental to the existing NERSC staff. One of these would be an HPSS developer charged with the understanding of HENP data access issues. LBNL(NERSC) is the only HENP site that is also an HPSS development site and hence the only site that can deploy a developer.

To ensure that there is no contention for core server resources, the ATLAS HPSS system must operate in its own HPSS namespace. The cost of additional namespaces at an HPSS deployment(user) site is about $250,000 plus $150,000 per year. HPSS development sites may deploy as many independent namespaces as desired without incurring the additional licensing costs. If the final deployment site(s) do not already have ongoing HPSS support, further costs are likely to be incurred to achieve a critical number of HPSS-support personnel. Our cost estimate, which is based on locating the ATLAS HPSS system at NERSC, represents a lower bound on this cost, and we estimate that the HPSS system support cost could be up to $1M per year higher at another site or sites.

The costs in this phase are dominated by the manpower. The costs associated with system support might be reduced if we used a shared facility such as the future NERSC-4. However, in this event, additional costs associated with code porting and maintenance might be incurred unless the shared facility were an ATLAS supported platform.

## 4.2 User and Developer Support

Support for users and developers is vital at every stage of the project. In addition to providing support for use of the central facility once established, U.S.-ATLAS users will expect assistance in exporting ATLAS code to their home institutions. ATLAS intends to limit the number of supported platforms. We have assumed, in our planning, support for two platforms.

U.S.-ATLAS will have a significant influence on the choice of supported platforms since it comprises 20% of the entire ATLAS effort. Software licenses for supported platforms will be negotiated, purchased, and distributed collaboration-wide by ATLAS. If U.S.-ATLAS concludes that it is economically advantageous to run hardware that is not supported by ATLAS, it would have to port the code, maintain it and negotiate all licenses locally.

Before effective code support and distribution at an ATLAS regional computing center can be implemented, a framework under which the ATLAS offline software is to be developed, distributed and maintained, must exist. Such a framework is being defined, but has not yet reached maturity. This makes it difficult now to make a proper assessment of the resources needed.

The size and complexity of the ATLAS experiment sets the scale of the requirements for the development, distribution, and maintenance of the offline source code repository, as well as the use of the offline software for analysis by the ATLAS members. Some problems are:

- The geographic separation of the developers and users of the offline software and the need to rely on the Internet and Internet-related tools to develop and distribute it.
- The diversity of the computing platforms that are located at the various development sites. The ability to support a heterogeneous computing environment is key to being able to take advantage of the latest computing technology.
- The large number of users of the offline software with widely varying skill levels.

A regional center will play the role of a mirror repository for its region. The master source code repository will reside on the CERN ATLAS computing system. Copies of the source code will migrate to the regional centers via wide-area network connectivity tools. Local users will be able to access a complete copy of the official ATLAS offline software and support software.

The Internet, via the ESNET backbone for the U.S. regional center, will be the primary means of delivering all the analysis software and related analysis tools provided by the ATLAS offline computing group.

The mechanism to carry out the distribution of the offline software source code repository will depend on the distribution method that ATLAS adopts. Currently, ATLAS/CERN is relying on AFS for their WAN needs. AFS, coupled with the current version of the Software Release Tools (SRT), provides the functionality of duplicating the CVS repository, which will be the first method used at the US ATLAS regional center. There are however a number of other possibilities for wide area software distribution tools that might be available in the future.

One possible role for the regional computing center would be to work with the ATLAS computing group to update SRT to accommodate the necessary mirroring software tools and then to use the regional center as a testbed for the new mirroring software. Whatever the form of the software distribution tools, the regional center will play an important role in developing and maintaining such software. The regional center may play a role in offloading work related to offline software maintenance from the CERN ATLAS offline group.

The regional center staff must also provide individual user support. As an example, NERSC provides a guaranteed level of service when a user contacts either User Services or Operations (after hours) which includes the following metrics:

- Respond to user problems within 4 working hours.
- Resolve 90% of user problems within 2 working days - Escalate within 72 hours if not resolved.
- All changes/outages announced at least 24 hours in advance.

- All planned changes/outages announced at least 7 days in advance.

The quality of service standards that we propose are discussed in Appendix A.

### 4.2.1 Development Phase

In the development phase, we anticipate that 2 FTEs will be required to provide infrastructure support for software developers and physicist developers and users.

Tasks performed by the two user support personnel include:

- Setting up and maintaining the environment (e.g. SRT/CVS repository) and tools (e.g. CASE tools, Insure++) necessary for the development of new ATLAS code.
- Database maintenance.
- Importing and distributing ATLAS code for simulation.
- Responding to routine user inquiries and problem reports.
- Setting up, authoring, and maintaining the U.S.-ATLAS Web pages.
- Authoring and organizing documentation and training.

Although there will certainly be some overlap in the responsibilities of the two support FTEs, we identify them as:

- Code Librarian: Primarily responsible for the first, second and third of the above tasks.
- Web Master & Documentarian: Primarily responsible for the remainder of the above tasks.

### 4.2.2 Deployment Phase

In the deployment phase, the support personnel will need to be ramped up from the development phase level to the analysis phase level. In preparation for the Mock Data Challenge, we will need to bring on a general user support person to help make ATLAS software accessible and usable by U.S.-ATLAS collaborators.

As the complexity of ATLAS software increases and more people contribute to the code, the importance of code verification and regression testing will become increasingly important. We expect that we will need a second librarian specifically charged with this task.

### 4.2.3 Analysis Phase

In the analysis phase, the number of users of the U.S.-ATLAS Regional Computing Center will be dramatically higher than during the development phase. The level of computer expertise and familiarity with ATLAS software of these users will cover a much wider range. While there will still be developers intimately familiar with the software, there will also be users familiar only with the user interface of the system, and largely ignorant of the underlying software architecture and components.

Addressing both the larger number of users and the larger range of questions and expertise will require five FTEs devoted to user support. The breakdown of effort will look like:

- 2 FTE Librarian/Code Verification/Regression Testing
  Responsible for maintaining the ATLAS software environment and developer tools, importing CERN software to the ARC, exporting ARC software to CERN, and distributing ATLAS software to U.S.-ATLAS institutions and collaborators. Also responsible for testing and verifying ATLAS code releases.
- 1 FTE Web/Training/Documentation
  Responsible for maintaining and coordinating Web-based documentation on U.S.-ATLAS and U.S.-ATLAS Computing, authoring and maintaining user documentation, and coordinating user training.
- 2 FTE General User Support
  Responsible for answering general user questions, assisting users with problems in understanding and using ATLAS software, and managing simulation production.

## 4.3 Training

It is critical that U.S.-ATLAS begin an aggressive program of training to assist physicists in making the transition from traditional programming methods to the Object-Oriented paradigm and to the use of Object-Oriented Databases. We will develop a training program, in close collaboration with the ATLAS management at CERN, that is tuned to the needs of the physicists working in areas managed by our software development function.

As an example, we cite the experience of the BaBar training program. In that program, every collaborator in the experiment who was involved in programming was encouraged to take an Intro-to-C++ course. These were held in conjunction with the regular collaboration meetings, and at other times. Approximately three hundred members of the BaBar experiment attended these.

All developers went to two commercial OO Design courses. These were four or five days in duration and had fifteen or fewer students, with a cost of approximately $1500 per student. The first course was an Introduction to OO Design with C++, and the second course was Advanced Design Techniques & Design Patterns. Approximately 100 collaborators have gone through both courses. Finally, 18 people went to a OODBMS course (commercial again) at a cost of $1000 per head and 4-5 days in duration.

We plan to initiate a similar program with training offered at a number of regional sites located near clusters of U.S.-ATLAS institutions (Boston, New York, Chicago and Berkeley). In addition, we will investigate computer-based and video-based training to reinforce material learned in the formal classes and will encourage video conferences among recent course graduates to share experiences as they begin to apply what they have learned.

As these courses are not inexpensive, we intend that they be budgeted centrally. This is essential if we are to encourage full involvement by U.S.-ATLAS members. We expect the cost of training in the development and deployment phases to be proportionately higher per year than in the analysis phase, because during analysis there will be many more U.S.-ATLAS collaborators familiar with OO techniques and with the software and because a higher percentage of new collaborators will be joining with prior OO experience and knowledge.

## 4.4 Networking

U.S.-ATLAS requires very high-bandwidth networking to all the laboratory and university sites. The network must support collaborative interactive data analysis as well as the transactions across the ATLAS federated database. In addition to bandwidth requirements, the network must provide low latency for some services such as database locking and interactive connections.

The Energy Sciences Network, ESnet, provides the backbone for all U.S. laboratory connections. Currently Berkeley and Argonne are connected to ESnet at OC12 (622 Mbits/sec.) The Brookhaven connection is now 45Mbits/sec but will likely be upgraded in early FY2000 after the new ESnet contract is completed.

We assume that all ATLAS universities will connect to each other and to the laboratories through Internet2, a consortium of research universities, or through another internet service provider. We expect that all university connections will move toward a minimum of OC3 (155 Mbits/sec.) before 2005. ESnet already has peering relationships with most networks and is working with Internet2 to establish peering with their new backbone network, Abeline. In our budget estimates, we assume that all network costs will continue to be funded by either the DOE, through ESnet, or by the universities as fees to Internet2 or an ISP. We have not included any explicit funding for networking to support U.S. ATLAS.

Networking between the United States and Europe, and CERN in particular, is significantly more expensive than inside the U.S. and for that reason, the available trans-Atlantic bandwidth has lagged that available in the U.S. This situation is changing as telecommunications in Europe are deregulated and as new carriers install trans-Atlantic bandwidth. It is clear that the bandwidth available between CERN and U.S.-ATLAS will go up as we approach the LHC turn-on, but it is not yet known if the capacity will be sufficient to permit transfer of ATLAS data. If not, we assume that we will transfer storage media, either tapes or some other technology, directly to the regional center for use by the U.S.-ATLAS physicists.

## 4.5 Near-Term Deliverables

Our near-term goals for infrastructure support are given in Table 6

| Date | Milestone |
|---|---|
| 5 Feb 1999 | Deploy Website for Software Projects (DONE) |
| 24 Feb 1999 | Deploy LHC++/ATLAS licenses on PDSF (DONE) |
| 1 Mar 1999 | Install SRT/CVS on PDSF (DONE) |
| Jun 1999 | GEANT4 training course; General U.S. Computing Workshop |
| Aug 1999 | Deploy disk and HPSS space for test beam data |
| | Assess training needs of U.S.-ATLAS |
| 30 Sept 1999 | Move test beam data from CERN |
| 30 Dec 1999 | Assess automatic code update systems |
| 30 Jan 2000 | First software training course in U.S. for developers |
| 30 Mar 2000 | Deploy code update system |
| | Web-based software documentation |

Table 6: Near-term milestones for infrastructure for U.S.-ATLAS Computing

### 4.6 Longer-Term Milestones

The key milestones for the partial and full deployment phases are given in Table 7:

| Date | Milestone |
|---|---|
| June 2000 | Deploy system for code export |
| | First software course for U.S.-ATLAS members (non-developers) |
| June 2001 | Begin process to determine 10% system for Mock-Data Challenge (MDC) |
| May 2002 | Decide on hardware purchase for 10% system |
| October 2002 | Decide method of data transport from CERN for MDC |
| March 2003 | Deploy 10% system |
| June 2003 | MDC begins |
| December 2003 | Begin acquisition process for full deployment |
| January 2004 | First MDC complete |
| June 2004 | Begin hardware acquisition for full system |
| June 2005 | Data taking begins |

Table 7: Key milestones for the partial deployment and full deployment phases

## 5 Cost and Resource Summary

The estimated costs for U.S.-ATLAS computing infrastructure and software development are summarized in Table 8 and the cost profile is shown in Figure 1. We emphasize that no physicist costs are included but that we expect that five additional postdoctoral fellows will be needed in FY2000 and ten in FY2001.

We have included a software engineer and clerical support as management costs even though this proposal does not include a management plan. The more detailed cost estimates for each of the major categories are provided in Appendix C. The costs of Management and Software Development include fully burdened personnel costs. Travel is included in "other costs." The estimates shown in Table 8 assume shared HPSS development and support.

We have taken a 50% contingency on hardware purchases starting in FY2001, a 50% contingency in user-support personnel, also starting in FY2001, and a 20% contingency in operations personnel. No contingency has been allocated to management or software development. Until the scope and deliverables of the software development by U.S.-ATLAS have been fixed, it makes little sense to allocate contingency in this case.

The costs given for personnel do not take into account possibilities for offsets from ongoing support. The amount of such offsets depends on the site(s) chosen for the regional center, but might amount to up to about 20% of the cost. The total cost with contingency up to and including FY2004 is 40.3 $M.

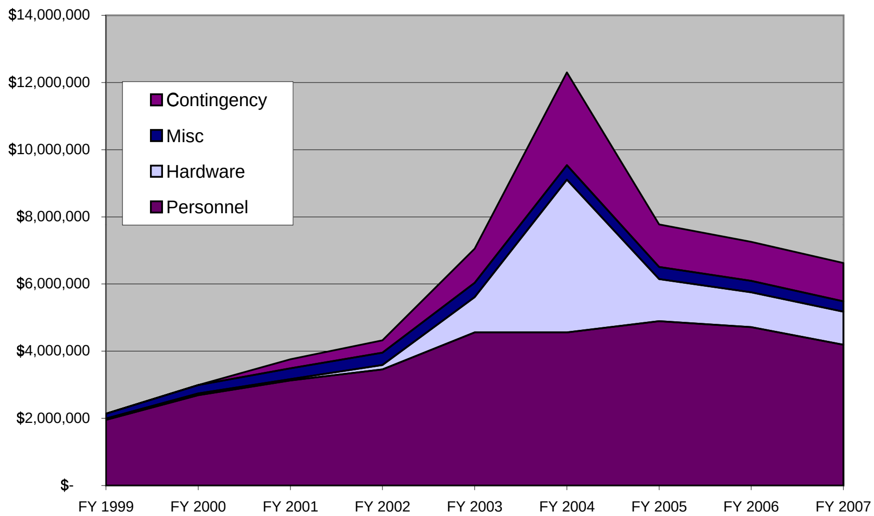


Figure 1: Cost profile.

| Fiscal Year | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 |
|---|---|---|---|---|---|---|---|---|---|
| Management | 1.5 FTE | 2.0 FTE | 2.0 FTE | 2.0 FTE | 2.0 FTE | 2.0 FTE | 2.0 FTE | 2.0 FTE | 2.0 FTE |
| ARC Personnel | 2.5 FTE | 3.9 FTE | 5.4 FTE | 6.4 FTE | 11.0 FTE | 11.0 FTE | 14.5 FTE | 14.5 FTE | 14.5 FTE |
| Code Developers | 7.5 FTE | 10 FTE | 11 FTE | 12 FTE | 14 FTE | 14 FTE | 13 FTE | 12 FTE | 9 FTE |
| Personnel cost | 1953346 | 2689042 | 3129131 | 3457821 | 4558080 | 4558080 | 4893058 | 4718058 | 4173058 |
| ARC Hardware | 20096 | 21883 | 9055 | 95917 | 1007778 | 4507858 | 1217633 | 996208 | 957999 |
| Developer Wrkstns | 22500 | 30000 | 33000 | 36000 | 42000 | 42000 | 39000 | 36000 | 27000 |
| Misc Costs | 143500 | 250500 | 317750 | 367000 | 429000 | 429000 | 355500 | 342500 | 303500 |
| Total Costs | 2139142 | 2991425 | 3488936 | 3956738 | 6036858 | 9536938 | 6515191 | 6092766 | 5481557 |
| Contingency. | 0 | 0 | 4527 | 47958 | 503889 | 2253929 | 608816 | 498104 | 478999 |
| Total with Contin. | 2139142 | 2991425 | 3758789 | 4324575 | 7053490 | 12303609 | 7776375 | 7253239 | 6622925 |

Table 8: U.S.-ATLAS Computing Cost Summary; all numbers in FY99 dollars.

# References


[1] The LHC Conceptual Design Report, http://www.cern.ch/CERN/LHC/YellowBook95/yellow.shtml

[2] ATLAS Computing Technical Proposal, CERN/LHCC 96 - 43 December 19, 1996 http://atlasinfor.cern.ch/Atlas/GROUPS/SOFTWARE/TDR/html Welcome.html

[3] Report of the ATLAS Computing Review Committee. http://edmsoraweb.cern.ch:8001/cedar/doc.download?document_id=1015

[4] Jurgen Knobloch http://www.cern.ch/Atlas/GROUPS SOFTWARE/INFO/Workshops/9812/ slides/tue.1/sld005.htm

[5] U.S.-ATLAS Green Book

[6] http://130.64.8.4/ATLAS/MONARC.html

[7] ATLAS Computing Milestones
http://www.cern.ch/Atlas/GROUPS/
SOFTWARE/milestones.html

[8] http://hepunx.rl.ac.uk/atlas/control

[9] http://home.cern.ch/ lat/exports/thesis

[10] Gordon Moore, Electronics Magazine, 1965. see Computers in Physics, 12, p 518, 1998.

[11] F.S. Preston, 1997

[12] Richard Mount in Proceedings of the International conference on Computing in High Energy and Nuclear Physics (CHEP98), Chicago IL, 31 August - 4 September, 1998.

[13] http://www.NERSC.gov

[14] http://www.nersc.gov/hardware/PDSF

[15] http://www.gridforum.org

[16] The Grid: Blueprint for a New Computing Infrastructure, I. Foster and C. Kesselman (Eds), Morgan-Kaufmann, 1999.

[17] An Architecture for Optimizing Query Processing and Data Delivery in Multilateral storage Environments, David M. Malon, in Proceedings of the International Conference on Computing in High Energy and Nuclear Physics, Chicago, - 31 August to 4 September, 1998
http://www-rnc.lbl.gov/GC/default.htm

[18] http://www.globus.org

[19] http://www-itg.lbl.gov/Clipper/

[20] The Particle Physics Data Grid, R. Mount *et.al.*

[21] http://now.cs.Berkeley.edu/Millennium

# A User Support and Operations Requirements

```
-- Analysis Phase

Hardware:

* CPU = 50k SPECin95
        - Average available CPU power over year.
* Disk = 25 TB
        - Minimum available Disk space.
* HPSS = 250 TB/year
        - Average increase in HPSS storage.

Operations Coverage:
* 24 X 7 coverage for monitoring all systems.

Coverage for monitoring the health & security of all systems should be
24 hours per day by 7 days per week. This implies operator-level
personnel on duty at all times and semi-automated software monitoring
tools and policies.
```

Monitored Systems include:

- CPU Nodes
- Central Servers
- Other Servers
- User Disks
- HPSS
- Data Disks
- Network

QoS - 95% of all failures and 75% of all performance degradations should be discovered through the operations/monitoring channel. (N.B. This applies only to system hardware and system level software, not to user code and/or batch jobs.).

* 24 X 7 coverage for recovering critical systems.

Coverage for recovering failures of critical systems should be 24 hours per day by 7 days per week. This implies sysadmin-level personnel on call during non-business hours.

Critical systems are those which are necessary for all analysis and simulation work. i.e. Absence of any critical system will make use of the ARC for analysis or simulation impossible.

N.B. We define a loss of capacity >= 40% in any non-critical system as a critical failure to be handled as the loss of a critical system.

Critical Systems include:

- Central Servers
- User Disks
- HPSS
- Data Disks
- Network

QoS of critical systems should be:

```
        - 95%           - Scheduled Availability
        - 90%           - Gross Availability
        - 336 h         - Mean Time Between Interruptions
        - 8 h           - Mean Time to Restoration
```

* 8 X 5 coverage for recovering non-critical systems.

Coverage for recovering failures of non-critical systems should be 8 hours per day by 5 days per week (Mon-Fri 9:00am-5:00pm excluding LBL holidays). This implies sysadmin-level personnel on duty during business hours.

Non-critical systems whose failure reduces capacity, but does not halt production and/or development work. i.e. The system as a whole can still be used even if some (but not all) non-critical systems have failed. The net result of such failures is a reduction in performance of the system (capacity) not a loss of functionality.

N.B. We define a loss of capacity >= 40% in any non-critical system as a critical failure to be handled as the loss of a critical system.

Non-Critical Systems include:

- CPU Nodes
- Other Servers
- Data Disks

QoS of non-critical systems should be:

- 98% - Scheduled Average Capacity
- 95% - Gross Average Capacity

User Support Coverage:

* 24 X 7 coverage for receiving user emergency reports.

Coverage for receiving user emergency reports should be 24 hours per day by 7 days per week. This implies a Web-based Emergency Report Form tied to the operator's pager and/or use of the NERSC 800 number.

Emergencies are conditions within the PDSF domain which affect users and/or facilities outside of the PDSF domain or conditions which, for every day they exist will cause significantly more than one day of detrimental effect or any problem which has the potential of irrevocable damage (including hardware damage and/or data loss).

QoS of user emergency report submission mechanism should be:

- 100% - Scheduled Availability
- 99% - Gross Availability
- 1440 h - Mean Time Between Interruptions
- 8 h - Mean Time to Restoration

* 24 X 7 coverage for receiving user support requests.

Coverage for receiving user support requests should be 24 hours per day by 7 days per week. This implies a Web-based support request Form and/or use of the NERSC 800 number.

QoS of user support request submission mechanism should be:

- 100% - Scheduled Availability
- 98% - Gross Availability
- 1440 h - Mean Time Between Interruptions
- 8 h - Mean Time to Restoration

* 24 X 7 coverage for responding to user emergency reports.

Coverage for responding to user emergency reports should be 24 hours per day by 7 days per week. This implies sysadmin-level personnel on call during non-business hours.

QoS of user emergency report response should be:

- 1 h - Mean Time to Response
- 4 h - Maximum Time to Response

* 8 X 5 coverage for responding to user support requests.

Coverage for responding to user support requests should be 8 hours per

```
day by 5 days per week (Mon-Fri 9:00am-5:00pm excluding LBL holidays).
This implies user-support personnel on duty during business hours.

QoS of user support requests response should be:
        - 8 hr          - Maximum Time to Response
        - 90%/40 hr     - 90% Tickets resolved within 40 business hours
```

# B Expertise in HEP Computing in the U.S-ATLAS

Physicists and software engineers at in U.S.-ATLAS have accumulated significant expertise that is directly relevant to U.S.-ATLAS Computing. They hold leadership roles in the computing systems for other major HEP and Nuclear Physics experiments including BaBar, CDF D0 and STAR, as well as contributing considerably to their overall computing effort. It is intended that this expertise will transition to U.S.-ATLAS Computing during the period where ATLAS software is being developed.

Physicists and Software Engineers from LBNL currently hold, or have held, positions of leadership responsibility within the computing systems of several HEP experiments. In particular, M. Shapiro is currently head of the CDF Off-line computing with overall responsibility for the off-line software for Run II. D. Quarrie was formerly Project Software Engineer for BaBar, with responsibility for establishing a software development infrastructure, training, budget and schedule. S. Patton was former Software Coordinator for CLEO. R. Jacobsen is off-line coordinator for BaBar.

Both ANL and LBNL were heavily involved with the PASS (Petabyte Access Store Solutions) project for the SSCL. Both Labs provided the management leadership (L. Price, S. Loken) and technical leadership (D. Quarrie), as well as several other members of that collaboration (C. Day (LBNL), D. Malon and E. May (ANL)). The PASS project broke new ground within HEP by proposing the use of a distributed object oriented environment based upon CORBA and ODMG-compliant object oriented database management system. It performed several studies that validated the use of these concepts, but was terminated before a prototype was developed.

Physicists and Software Engineers from LBNL have designed and implemented the control domain environments for several large HEP experiments. M. Shapiro and D. Quarrie wrote the original such framework for CDF using FORTRAN, and D. Quarrie implemented a successor to this using C++ and OO techniques. This is now used by both BaBar and CDF. S. Patton was co-author of the equivalent C++ based framework that is in use by CLEO-III. C. Tull is the author of the analysis framework in use by STAR.

Physicists from BNL currently hold, or have held, positions of leadership responsibility within the computing systems of several HEP and NP experiments. In particular, S. Protopopescu is currently head of the D0 Run II Algorithms Group and was head of all of D0 off-line software for Run 1, T. Wenaus is the head of STAR off-line computing, S. Adler was for many years the D0 software librarian and has since become an expert in Unix administration as well as physics analysis, S. Snyder is an expert in data persistence, S. Rajagopolan was a major player in the D0 online and offline effort and responsible for Run I database and tracking algorithms, T. Satogata is an expert in real time applications for the RHIC machine and OO code development, N. Graf is co-head of the D0 Run II global tracking.

Physicists and software engineers from ANL (E. May and D. Malon) hold positions with significant design and implementation responsibility within the HENP Grand Challenge Project. This project utilizes a distributed CORBA based architecture for high performance physics analysis. It has pioneered the use of database indexing techniques as an alternative to conventional iteration. The initial implementation is in use by the STAR experiment, and similar indexing techniques have been adopted by the BaBar experiment.

Overall responsibility for databases for the BaBar experiment is provided by a LBNL group lead by D. Quarrie. S. Patton is senior engineer for the event store, including data distribution to regional centers, I. Gaponenko is in charge of the design and implementation for the calibrations/alignment database, and C. Day is responsible for the integration with the online. The calibration/alignment database design has been adopted by the RD45 collaboration as the basis for LHC experiments.

Participants in the U.S. software effort have been members of the RD45 project at CERN since it's inception and have provided significant contributions to its prototypes (e.g. calibration/alignment database). E. May and D. Malon (ANL), D. Quarrie (LBNL), Krzys Sliwa and Simone Rolli (Tufts U.).

There is significant activity in the U.S. University groups on ATLAS software in addition to that discussed in detail elsewhere in the proposal. Persons include; Fred Luehring (Indiana U.) who is software coordinator for the TRT and is responsible for simulation of the TRT; Bill Seligman and Misha Leltchouk (Columbia U.) are involved in reconstruction and simulation of the LAr calorimeter; Toby Burnett (U. Washington) was the author of ARVE, the current ATLAS framework. Krzys Sliwa (Tufts U.) is chair of ATLAS World-Wide Computing Group and was instrumental in the creation of MONARC. Steve Goldfarb's (Michigan), is working on ATLAS database architecture and has served as the scientific secretary for the weekly ATLAS worldwide Software Meetings. Bob Jacobsen and Homer Neal served on the recent ATLAS computing review.

In addition many members of U.S.-ATLAS have been involved intensively in the work for the physics TDR and have worked on software as part of this effort. In particular groups in Boston, Michigan, and Columbia have carried out detailed simulation studies.

In addition, considerable expertise in object oriented techniques such as C++, UML, Java, OODBMS is available, as well as creation of an effective software development infrastructure. Within the LBNL group M. Marino provides CASE expertise within the BaBar experiment, critically examining existing software designs, testing them against quality metrics and providing constructive feedback to the original designers. A. Mohktarani provides general reconstruction support for BaBar, in particular in support of the database, focussed on physics usage.

## B.1 Computing Expertise of the Laboratories and Universities.

The National Energy Research Scientific Computing Center (NERSC) [13] is a national facility that provides very large-scale computing and storage facilities to programs funded by the DOE Office of Science (DOE/SC). This facility was founded in 1974 at LLNL and was moved to LBNL in 1996. This move was intended to significantly improve the communication for NERSC with its users and with the computer science community who are producing the tools and frameworks for the future of high-end computing and storage. Following the move to Berkeley, the management of NERSC committed to the support of the PDSF [14] for the High Energy and Nuclear Physics communities in DOE/SC. This includes operation of computers (now PCs and Sun servers) and a high-speed connection to the NERSC mass storage facility which is based on the High-Performance Storage System (HPSS).

HPSS (a hierarchical storage system (disk and tape) based on the IEEE Mass Storage Standard) is a complex, highly configurable system. It allows customization to the specific access patterns by users of the mass storage system. NERSC has been a member of the HPSS collaboration since its inception by IBM and five DOE national laboratories. It is one of five development sites. The NERSC user and backup HPSS systems routinely handle 300 GB per day of new data each as well as associated reads of data. SLAC computer scientists performed systematic studies of HPSS read and write performance as a function of block size, quantitatively demonstrating the motivation for using a system like Andy Hanushevsky's OOFS when accessing Objectivity database files stored on HPSS.

Being an HPSS development site has many advantages for the ATLAS Regional Computing Center. For example, HPSS development sites are permitted to run an unlimited number of independent production HPSS systems without charge. Under the current HPSS licensing scheme, each additional HPSS namespace (i.e. each distinct system of core servers and independent hardware.) put into production at a customer site requires an additional license purchase fee of $300,000 and an additional maintenance fee of $150,000 per year. It is not clear whether a separate ATLAS HPSS system, where there is no contention with non-ATLAS users for HPSS core services, is preferable to one where ATLAS-specific HPSS Classes Of Service are offered in the shared HPSS systems. In addition, HPSS development sites are legally permitted to run local modifications to the HPSS code. This would enable the ATLAS HPSS system at NERSC to be customized beyond what may be possible with operational parameters at non development sites.

The regional center will have to import data into HPSS. NERSC has twice performed imports of large quantities of data. In November 1997, the NERSC MSG converted the NERSC Unitree system containing more than 500,000 files and 10 TB of user data to HPSS. In January 1999, the NERSC CFS system containing more than 20,000 tape cartridges and 20 TB of user data was converted to HPSS. We are confident that data can be moved from CERN.

LBNL is the host laboratory for the Energy Sciences Networking group and ESnet Network Operations Center. This provides maximum connectivity between the ARC and other ATLAS institutions (i.e. collaborating national labs and U.S. universities). Argonne has a very fast connection to ESnet, plus an unrivaled set of connections to national and international research networks that permit Argonne and Berkeley to provide the best possible support for code development and data analysis at all U.S.-ATLAS institutions. The connection from Argonne to ESnet is presently at OC-12 speed (622 Mbs) over ATM and will support very tight integration of the U.S.-ATLAS computing work between Argonne and Berkeley. Argonne has led an extensive set of Midwestern universities in setting up an MREN (Metropolitan Research and Education Network) connection at OC-3 (155 Mbs) and OC-12 speeds and is connected to the NSF vBNS at OC-12 speed, providing excellent connectivity to university collaborators. In the international arena, Argonne is an operator of the NSF STAR-TAP, which is becoming the U.S. interconnect point for international networking across both the Atlantic and Pacific Oceans. Within ESnet, Argonne has provided leadership in international planning and collaborations.

The RHIC Computing Facility (RCF) is an HENP specific computing facility established to satisfy the offline computing needs of the RHIC experiments at Brookhaven National Laboratory. In terms of data volume handled and processing power required the RHIC experiments need computing capacities in excess of those previously required. The data volume produced at RHIC will be in excess of a PByte per year. Tape transfer rates will be of order 300 MBytes/sec, while disk transfer rates will be in excess of a GByte/sec. The amount of installed CPU required for reconstruction, data mining and analysis is approximately 20,000 SPECint95. The four RHIC experimental collaborations include nearly 1000 physicists from 80 institutions in 20 countries. Since the qualitative computing requirements and the appropriately normalized scale are nearly the same, the technical solutions being developed and deployed at RHIC are similar to those propose for the LHC experiments.

The supported computing platforms at RCF are SUN/Solaris and Intel/Linux. The farms, depending on mission, are managed either by the LSF resource management (batch) system or by in house developed reconstruction management software. AFS and Objectivity are also deployed on these Linux systems, RCF having performed the initial port of Objectivity to Linux.

For file level data storage and management, RCF utilizes the High Performance Storage System (HPSS) which is in use or under test by several other HENP sites including CERN. HPSS required special interfacing to produce acceptable HEP behavior. To this end RCF has integrated selected third party interface components from ORNL and CERN and has also develop additional HPSS access software to manage casual user archiving and access by the Linux reconstruction farm.

Over the course of the last six months RCF and the experimental groups have engaged in two Mock Data exercises at the 10-15% of design capacity level. These have validated primary technical choices and allowed for iterative adjustment of technical and operational details. These exercises have included the OO designed and developed experiment application codes and the operation of data access tools being developed for RHIC use by a DOE Mathematics, Information and Computational Science (MICS) funded Grand Challenge Project. Within three months the first samples of actual experimental data are expected. Sustained processing of physics data will begin in November 1999 and the facility will be in full operation at design levels with 24 months

## B.2 Community Software Projects.

In addition to the ATLAS Computing activities that are carried out by the collaboration, U.S.-ATLAS efforts will benefit from a number of activities in the Computer Science community and from collaborations with other experiments. One such joint project is MONARC, a collaboration with CERN and CMS which is discussed elsewhere. Members of U.S.-ATLAS are also involved in the work on the High Energy and Nuclear Physics Grand Challenge[17] Project, the Globus Middleware Project [18], the Clipper Project[19], Next Generation Internet (NGI)[20], the Science Simulation Initiative (SSI) and the DOE2000 Collaboratory Projects.

The HENP Grand Challenge [17] has been a very successful collaboration of physicists and computer scientists aimed at creating a prototype data handling architecture for experiments at RHIC and LHC. The project has taken advantage of extensive expertise in data management at Argonne and at Berkeley. Both Argonne and Berkeley have been leaders in research and development of scalable technologies for particle physics data storage and access. Argonne has contributed significantly to the overall Grand Challenge architectural design, and has taken responsibility, in particular, for client-side components. In the Grand Challenge, LBNL developed the technology and software to manage Terabytes of HENP data and their movement to a shared disk cache.

The Argonne/ISI Globus project [18] has developed a range of basic services designed to support efficient and effective computation in Grid environments. These services have been deployed in a large multi-institutional testbed that spans some 40 sites worldwide and has been used in numerous application projects ranging from collaborative design to distributed supercomputing. Services on the Grid include

- The Grid Security Infrastructure (GSI), which provides public key-based single sign-on, run anywhere capabilities for multi-site environments, supporting proxy credentials, interoperability with local security mechanisms, local control over access, and delegation.
- The Metacomputing Directory Service (MDS), which provides a uniform representation of, and access to, information about the structure and state of Grid resources, including computers, networks, and software.
- Globus resource management services, which provide uniform resource allocation, object creation, computation management, and co-allocation mechanisms for diverse resource types. The Global Access to Secondary Storage (GASS) service, which provides a uniform name space (via URLs) and access mechanisms for files accessed via different protocols and stored in diverse storage system types (HTTP, FTP are currently supported; HPSS, DPSS are under development.

The Clipper Project is a joint project with ANL, LBNL, and SLAC which is focused on developing technologies required for widely distributed data-intensive applications, in particular particle-physics data analysis. Clipper leverages existing technologies such as the Distributed Parallel Storage System (DPSS), Globus, and ESnet and NTON (OC-12 networks) with the goal of achieving high-performance, guaranteed data-rate bulk transfers. In one set of initial experiments, data transfer rates between LBNL and SLAC for a particle-physics application. equivalent to moving 4.5 terabytes/day were demonstrated. In other experiments, data rates of 320 megabit/s were demonstrated across an OC12 ESnet link between ANL and LBNL. This work has demonstrated the feasibility of the data transfer rates required for U.S.-ATLAS and has also provided valuable experience in the network instrumentation, optimization, and debugging required to achieve high sustained rates in wide area networks.

Members of U.S.-ATLAS are also participants in a proposal to DOE's Next Generation Internet program. The Particle Physics Data Grid has two objectives: delivery of an infrastructure for widely distributed analysis of particle physics data at multi-Petabyte scales by thousands of physicists, and acceleration of the development of network and middleware infrastructure aimed broadly at data-intensive collaborative science. The project will design, develop, and deploy a network and middleware infrastructure capable of supporting data analysis and data flow patterns common to the many particle physics experiments represented among the collaboration.

Application-specific software will be adapted to operate in this wide-area environment and to exploit this infrastructure. The result of these collaborative efforts will be the instantiation and delivery of an operating infrastructure for distributed data analysis by participating physics experiments such as the U.S.-ATLAS TileCal Testbeam project.

The NGI [20]work will also be relevant to the DOE Science Simulation Initiative. Collaborative access to large data volumes will become a requirement for progress in many sciences, and especially the Global Systems and Combustion thrusts of SSI. These sciences have a data management and visualization problem very different from that of particle physics. However, we expect a convergence of the data-management needs of particle physics and many other sciences over the next decade.

Finally, a number of U.S.-ATLAS groups are working on collaboratory technologies which will significantly enhance our ability to work together within the U.S. and with our colleagues in ATLAS. in ways that would not be possible without high-speed networks. The Department of Energy, through its DOE2000 program, has initiated an effort to develop and enhance the tools needed to make all of the DOE facilities accessible to scientists in the United States and to improve the ability of those scientists to work together. Both Argonne and Berkeley are major contributers to the development of collaboratory tools and to their application in science projects.

Another important Collaboratory effort is being led by physicists and computer scientists at the University of Michigan.

Michigan has been a major contributor to networking and collaborative work. Through its close affiliation with Internet2, based in Ann Arbor, it has spearheaded discussions that have led to CERN having high speed network connections into STARTAP, and has helped prepare the groundwork for CERN joining UCAID and having access to ABILENE, steps that will clearly be of considerable value to U.S.-ATLAS computing. In addition, the University of Michigan School of Information has pioneered the use of collaborative tools for research in the atmospheric sciences, a field that faces many of the same issues of distributed research as high energy physics. The ATLAS group at Michigan has initiated a collaboration with CERN and CMS to further enhance the collaboration tools and to integrate them into ATLAS. This effort will be integrated with the DOE2000 efforts that are underway at Argonne and Berkeley.

# C Cost Details

# ARC Costs (FY99$)

| ARC Costs (FY99$) | October-98 | October-99 | October-00 | October-01 | October-02 | October-03 | October-04 | October-05 | October-06 |
|---|---|---|---|---|---|---|---|---|---|
| **ARC Costs (FY99$)** | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| Managment FTEs | 0.50 FTE | 0.90 FTE | 0.90 FTE | 0.90 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE |
| Support FTEs | 1.00 FTE | 2.00 FTE | 2.50 FTE | 3.00 FTE | 4.00 FTE | 4.00 FTE | 5.00 FTE | 5.00 FTE | 5.00 FTE |
| Operations FTEs | 1.00 FTE | 1.00 FTE | 2.00 FTE | 2.50 FTE | 6.00 FTE | 6.00 FTE | 8.50 FTE | 8.50 FTE | 8.50 FTE |
| *Total FTEs* | *2.50 FTE* | *3.90 FTE* | *5.40 FTE* | *6.40 FTE* | *11.00 FTE* | *11.00 FTE* | *14.50 FTE* | *14.50 FTE* | *14.50 FTE* |
| Personnel Costs (FY99$) | $ 427,876 | $ 683,042 | $ 948,131 | $ 1,101,821 | $ 1,852,080 | $ 1,852,080 | $ 2,362,058 | $ 2,362,058 | $ 2,362,058 |
| Personnel Contingency (FY99$) | $ - | $ - | $ 265,325 | $ 319,879 | $ 512,742 | $ 512,742 | $ 662,368 | $ 662,368 | $ 662,368 |
| *Personnel Funding Needed (FY99$)* | *$ 427,876* | *$ 683,042* | *$ 1,213,456* | *$ 1,421,700* | *$ 2,364,822* | *$ 2,364,822* | *$ 3,024,426* | *$ 3,024,426* | *$ 3,024,426* |
| Hardware Costs (FY99$) | $ 20,096 | $ 21,883 | $ 9,055 | $ 95,917 | $ 1,007,778 | $ 4,507,858 | $ 1,217,633 | $ 996,208 | $ 957,999 |
| Hardware Contingency (FY99$) | $ - | $ - | 4,527.54 | 47,958.35 | 503,889.19 | 2,253,929.18 | 608,816.45 | 498,104.19 | 478,999.54 |
| *Hardware Funding Needed (FY99$)* | *$ 20,096* | *$ 21,883* | *$ 13,583* | *$ 143,875* | *$ 1,511,668* | *$ 6,761,788* | *$ 1,826,449* | *$ 1,494,313* | *$ 1,436,999* |
| Other Costs (FY99$) | $ 18,000 | $ 96,500 | $ 150,750 | $ 187,000 | $ 223,000 | $ 223,000 | $ 162,500 | $ 162,500 | $ 162,500 |
| Other Contingency (FY99$) | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - |
| *Other Funding Needed (FY99$)* | *$ 18,000* | *$ 96,500* | *$ 150,750* | *$ 187,000* | *$ 223,000* | *$ 223,000* | *$ 162,500* | *$ 162,500* | *$ 162,500* |
| *Total Costs (FY99$)* | *$ 465,972* | *$ 801,425* | *$ 1,107,936* | *$ 1,384,738* | *$ 3,082,858* | *$ 6,582,938* | *$ 3,742,191* | *$ 3,520,766* | *$ 3,482,557* |
| *Total Contingency (FY99$)* | *$ -* | *$ -* | *$ 269,853* | *$ 367,837* | *$ 1,016,631* | *$ 2,766,671* | *$ 1,271,185* | *$ 1,160,473* | *$ 1,141,368* |
| ***Funding Needed (FY99$)*** | ***$ 465,972*** | ***$ 801,425*** | ***$ 1,377,789*** | ***$ 1,752,575*** | ***$ 4,099,490*** | ***$ 9,349,609*** | ***$ 5,013,375*** | ***$ 4,681,239*** | ***$ 4,623,925*** |
| | | | | | | | | | |
| **Computer Costs (FY99$)** | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| Personnel Costs (FY99$) | $ 148,611 | $ 148,611 | $ 148,611 | $ 222,916 | $ 525,217 | $ 525,217 | $ 876,424 | $ 876,424 | $ 876,424 |
| Hardware Costs (FY99$) | $ 20,096 | $ 21,883 | $ 9,055 | $ 95,917 | $ 613,732 | $ 4,136,264 | $ 745,273 | $ 712,149 | $ 690,067 |
| Other Costs (FY99$) | $ - | $ 12,000 | $ 30,000 | $ 30,000 | $ 60,000 | $ 60,000 | $ 60,000 | $ 60,000 | $ 60,000 |
| Total Costs (FY99$) | $ 168,707 | $ 170,494 | $ 157,666 | $ 318,833 | $ 1,138,949 | $ 4,661,481 | $ 1,621,697 | $ 1,588,574 | $ 1,566,491 |
| **Support Costs (FY99$)** | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| Personnel Costs (FY99$) | $ 158,770 | $ 317,540 | $ 396,925 | $ 476,310 | $ 635,080 | $ 635,080 | $ 793,850 | $ 793,850 | $ 793,850 |
| Hardware Costs (FY99$) | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - |
| Other Costs (FY99$) | $ - | $ 66,500 | $ 99,750 | $ 133,000 | $ 133,000 | $ 133,000 | $ 66,500 | $ 66,500 | $ 66,500 |
| Total Costs (FY99$) | $ 158,770 | $ 317,540 | $ 396,925 | $ 476,310 | $ 635,080 | $ 635,080 | $ 793,850 | $ 793,850 | $ 793,850 |
| **MSS Costs (FY99$)** | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| Personnel Costs (FY99$) | $ - | $ - | $ 185,704 | $ 185,704 | $ 450,793 | $ 450,793 | $ 450,793 | $ 450,793 | $ 450,793 |
| Hardware Costs (FY99$) | $ - | $ - | $ - | $ - | $ 394,046 | $ 371,594 | $ 472,360 | $ 284,059 | $ 267,932 |
| Other Costs (FY99$) | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - |
| Total Costs (FY99$) | $ - | $ - | $ 185,704 | $ 185,704 | $ 844,840 | $ 822,387 | $ 923,153 | $ 734,852 | $ 718,725 |
| **Management Costs (FY99$)** | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| Personnel Costs (FY99$) | $ 120,495 | $ 216,891 | $ 216,891 | $ 216,891 | $ 240,990 | $ 240,990 | $ 240,990 | $ 240,990 | $ 240,990 |
| Hardware Costs (FY99$) | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - |
| Other Costs (FY99$) | $ 18,000 | $ 18,000 | $ 21,000 | $ 24,000 | $ 30,000 | $ 30,000 | $ 36,000 | $ 36,000 | $ 36,000 |
| Total Costs (FY99$) | $ 138,495 | $ 234,891 | $ 237,891 | $ 240,891 | $ 270,990 | $ 270,990 | $ 276,990 | $ 276,990 | $ 276,990 |

| Computer Costs (FY99$) | | October-98 | October-99 | October-00 | October-01 | October-02 | October-03 | October-04 | October-05 | October-06 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Computer Costs (FY99$)** | | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| Personnel Costs (FY99$) | | $ 148,611 | $ 148,611 | $ 148,611 | $ 222,916 | $ 525,217 | $ 525,217 | $ 876,424 | $ 876,424 | $ 876,424 |
| Hardware Costs (FY99$) | | $ 20,096 | $ 21,883 | $ 9,055 | $ 95,917 | $ 613,732 | $ 4,136,264 | $ 745,273 | $ 712,149 | $ 690,067 |
| Other Costs (FY99$) | | $ - | $ 12,000 | $ 30,000 | $ 30,000 | $ 60,000 | $ 60,000 | $ 60,000 | $ 60,000 | $ 60,000 |
| Total Costs (FY99$) | | $ 168,707 | $ 170,494 | $ 157,666 | $ 318,833 | $ 1,138,949 | $ 4,661,481 | $ 1,621,697 | $ 1,588,574 | $ 1,566,491 |
| | | | | | | | | | | |
| **CPU Power Required** | | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| CPU Required | | 100 Si95 | 100 Si95 | 100 Si95 | 500 Si95 | 5000 Si95 | 50000 Si95 | 50000 Si95 | 50000 Si95 | 50000 Si95 |
| CPU Purchased | | 100.00 Si95 | 25.00 Si95 | 25.00 Si95 | 425.00 Si95 | 4625.00 Si95 | 46250.00 Si95 | 12500.00 Si95 | 12500.00 Si95 | 12500.00 Si95 |
| CPU Retired | | 0.00 Si95 | 25.00 Si95 | 25.00 Si95 | 25.00 Si95 | 125.00 Si95 | 1250.00 Si95 | 12500.00 Si95 | 12500.00 Si95 | 12500.00 Si95 |
| **CPU Cost (FY99$)** | | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| CPU Moore's Factor | 1.50 | 1.00 | 0.67 | 0.44 | 0.30 | 0.20 | 0.13 | 0.09 | 0.09 | 0.09 |
| Linux CPU Price (FY99$/Si95) | | $ 168.28 | $ 112.19 | $ 74.79 | $ 49.86 | $ 33.24 | $ 22.16 | $ 14.77 | $ 14.77 | $ 14.77 |
| Sun CPU Price (FY99$/Si95) | **(Percent)** | $ 1,008.88 | $ 672.59 | $ 448.39 | $ 298.93 | $ 199.28 | $ 132.86 | $ 88.57 | $ 88.57 | $ 88.57 |
| Linux CPU Cost (FY99$) | 50% | $ 16,827.89 | $ 1,402.32 | $ 934.88 | $ 10,595.34 | $ 76,868.12 | $ 512,454.11 | $ 92,334.07 | $ 92,334.07 | $ 92,334.07 |
| Sun CPU Cost (FY99$) | 50% | $ - | $ 8,407.32 | $ 5,604.88 | $ 63,521.96 | $ 460,845.56 | $ 3,072,303.76 | $ 553,568.25 | $ 553,568.25 | $ 553,568.25 |
| *CPU Cost (FY99$)* | 100% | *$ 16,827.89* | *$ 9,809.64* | *$ 6,539.76* | *$ 74,117.29* | *$ 537,713.68* | *$ 3,584,757.87* | *$ 645,902.32* | *$ 645,902.32* | *$ 645,902.32* |
| **Disk Space Required** | | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| Disk Required | | 50 GB | 250 GB | 250 GB | 1000 GB | 5000 GB | 50000 GB | 50000 GB | 50000 GB | 50000 GB |
| Disk Purchased | | 50 GB | 200 GB | 63 GB | 813 GB | 4250 GB | 46250 GB | 12500 GB | 12500 GB | 12500 GB |
| Disk Retired | | 0 GB | 0 GB | 63 GB | 63 GB | 250 GB | 1250 GB | 12500 GB | 12500 GB | 12500 GB |
| **Disk Cost (FY99$)** | | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| Disk Moore's Factor | 1.50 | 1.00 | 0.67 | 0.44 | 0.30 | 0.20 | 0.13 | 0.09 | 0.06 | 0.04 |
| Disk Vault Price (FY99$/GB) | | $ 65.36 | $ 43.57 | $ 29.05 | $ 19.37 | $ 12.91 | $ 8.61 | $ 5.74 | $ 3.83 | $ 2.55 |
| Sun Raid Price (FY99$/GB) | | $ 208.33 | $ 138.89 | $ 92.59 | $ 61.73 | $ 41.15 | $ 27.43 | $ 18.29 | $ 12.19 | $ 8.13 |
| Sun SCSI Price (FY99$/GB) | **(Percent)** | $ 92.59 | $ 61.73 | $ 41.15 | $ 27.43 | $ 18.29 | $ 12.19 | $ 8.13 | $ 5.42 | $ 3.61 |
| Disk Vault Cost (FY99$) | 50% | $ 3,268.10 | $ 4,357.47 | $ 907.81 | $ 7,867.66 | $ 27,435.93 | $ 199,044.96 | $ 35,863.96 | $ 23,909.30 | $ 15,939.54 |
| Sun Raid Cost (FY99$) | 10% | $ - | $ 2,777.78 | $ 578.70 | $ 5,015.43 | $ 17,489.71 | $ 126,886.15 | $ 22,862.37 | $ 15,241.58 | $ 10,161.05 |
| Sun SCSI Cost (FY99$) | 40% | $ - | $ 4,938.27 | $ 1,028.81 | $ 8,916.32 | $ 31,092.82 | $ 225,575.37 | $ 40,644.21 | $ 27,096.14 | $ 18,064.09 |
| *Disk Cost (FY99$)* | 100% | *$ 3,268.10* | *$ 12,073.52* | *$ 2,515.32* | *$ 21,799.41* | *$ 76,018.46* | *$ 551,506.48* | *$ 99,370.54* | *$ 66,247.02* | *$ 44,164.68* |
| **Level of Effort Required** | | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| ARC Operation | | 0.33 FTE | 0.33 FTE | 0.33 FTE | 0.50 FTE | 1.00 FTE | 1.00 FTE | 2.50 FTE | 2.50 FTE | 2.50 FTE |
| ARC Administration | | 0.67 FTE | 0.67 FTE | 0.67 FTE | 1.00 FTE | 2.50 FTE | 2.50 FTE | 3.50 FTE | 3.50 FTE | 3.50 FTE |
| ARC Architecture Porting | | 0.00 FTE | 0.00 FTE | 0.00 FTE | 0.00 FTE | 0.00 FTE | 0.00 FTE | 0.00 FTE | 0.00 FTE | 0.00 FTE |
| *Level of Effort Required* | | *1.00 FTE* | *1.00 FTE* | *1.00 FTE* | *1.50 FTE* | *3.50 FTE* | *3.50 FTE* | *6.00 FTE* | *6.00 FTE* | *6.00 FTE* |
| **Fully burdened FTE Cost** | **Base Cost** | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| ARC Operation | $ 128,292 | $ 42,764 | $ 42,764 | $ 42,764 | $ 64,146 | $ 128,292 | $ 128,292 | $ 320,730 | $ 320,730 | $ 320,730 |
| ARC Administration | $ 158,770 | $ 105,847 | $ 105,847 | $ 105,847 | $ 158,770 | $ 396,925 | $ 396,925 | $ 555,695 | $ 555,695 | $ 555,695 |
| ARC Architecture Porting | $ 158,770 | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - |
| *Fully burdened FTE Cost* | | *$ 148,611* | *$ 148,611* | *$ 148,611* | *$ 222,916* | *$ 525,217* | *$ 525,217* | *$ 876,424* | *$ 876,424* | *$ 876,424* |
| TestBed Costs (FY99$) | | $ - | $ 10,000 | $ 25,000 | $ 25,000 | $ 50,000 | $ 50,000 | $ 50,000 | $ 50,000 | $ 50,000 |
| Software Costs (FY99$) | | $ - | $ 2,000 | $ 5,000 | $ 5,000 | $ 10,000 | $ 10,000 | $ 10,000 | $ 10,000 | $ 10,000 |
| | | | | | | | | | | |
| **Supporting Information** | | | | | | | | | | |
| 400 MHz PII Linux (128 MB/11 GB) | | $ 2,575 | 15.30 Si95 | | | | | | | |
| 14 node Sun E4000 (1GB) | | $ 250,000 | 247.80 Si95 | | | | | | | |
| 400 MHz PII Linux (256 MB/64 GB) | | $ 4,131 | 63 GB | | | | | | | |
| 400 MHz Quad Sun E450 (1GB) | | $ 30,000 | | | | | | | | |
| 8X18 GB Sun Raid (24 MB cache) | | $ 25,000 | 144 GB | | | | | | | |
| 9X18 GB Vanguard Disk Unit (Ultra Wide SCSI) | | $ 10,000 | 162 GB | | | | | | | |
| Sun/Sun Raid System | | $ 180,000 | 864 GB | | | | | | | |
| Sun/Vangard Disk System | | $ 90,000 | 972 GB | | | | | | | |

| **Management Costs (FY99$)** | | **October-98** | **October-99** | **October-00** | **October-01** | **October-02** | **October-03** | **October-04** | **October-05** | **October-06** |
|---|---|---|---|---|---|---|---|---|---|---|
| **Management Costs (FY99$)** | | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| Personnel Costs (FY99$) | | $ 120,495 | $ 216,891 | $ 216,891 | $ 216,891 | $ 240,990 | $ 240,990 | $ 240,990 | $ 240,990 | $ 240,990 |
| Hardware Costs (FY99$) | | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - |
| Other Costs (FY99$) | | $ 18,000 | $ 18,000 | $ 21,000 | $ 24,000 | $ 30,000 | $ 30,000 | $ 36,000 | $ 36,000 | $ 36,000 |
| Total Costs (FY99$) | | $ 138,495 | $ 234,891 | $ 237,891 | $ 240,891 | $ 270,990 | $ 270,990 | $ 276,990 | $ 276,990 | $ 276,990 |
| | | | | | | | | | | |
| | | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| **Total FTEs** | | 2.50 FTE | 3.90 FTE | 5.40 FTE | 6.40 FTE | 11.00 FTE | 11.00 FTE | 14.50 FTE | 14.50 FTE | 14.50 FTE |
| **Level of Effort Required** | | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| ARC Facility Lead | | 0.50 FTE | 0.90 FTE | 0.90 FTE | 0.90 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE |
| *Level of Effort Required* | | *0.50 FTE* | *0.90 FTE* | *0.90 FTE* | *0.90 FTE* | *1.00 FTE* | *1.00 FTE* | *1.00 FTE* | *1.00 FTE* | *1.00 FTE* |
| **Fully Burdened FTE Cost (FY99$)** | **Base Cost** | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| ARC Facility Lead | $ 240,990 | $ 120,495 | $ 216,891 | $ 216,891 | $ 216,891 | $ 240,990 | $ 240,990 | $ 240,990 | $ 240,990 | $ 240,990 |
| *Fully Burdened FTE Cost (FY99$)* | | *$ 120,495* | *$ 216,891* | *$ 216,891* | *$ 216,891* | *$ 240,990* | *$ 240,990* | *$ 240,990* | *$ 240,990* | *$ 240,990* |
| **Miscellaneous Cost (FY99$)** | | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| Travel (# of trips to CERN) | **(Base Cost)** | 6 | 6 | 7 | 8 | 10 | 10 | 12 | 12 | 12 |
| Travel Cost (FY99$) | $ 3,000 | $ 18,000 | $ 18,000 | $ 21,000 | $ 24,000 | $ 30,000 | $ 30,000 | $ 36,000 | $ 36,000 | $ 36,000 |

## Management Costs (FY99$)

| | | October-98 | October-99 | October-00 | October-01 | October-02 | October-03 | October-04 | October-05 | October-06 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Management Costs (FY99$)** | | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| Personnel Costs (FY99$) | | $ 213,000 | $ 256,000 | $ 256,000 | $ 256,000 | $ 256,000 | $ 256,000 | $ 256,000 | $ 256,000 | $ 256,000 |
| Hardware Costs (FY99$) | | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - |
| Other Costs (FY99$) | | $ 28,000 | $ 24,000 | $ 24,000 | $ 24,000 | $ 24,000 | $ 24,000 | $ 24,000 | $ 24,000 | $ 24,000 |
| Total Costs (FY99$) | | $ 241,000 | $ 280,000 | $ 280,000 | $ 280,000 | $ 280,000 | $ 280,000 | $ 280,000 | $ 280,000 | $ 280,000 |
| Personnel Offset (FY99$) | | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - |
| Hardware Offset (FY99$) | | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - |
| Other Offsets (FY99$) | | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - |
| Total Offset (FY99$) | | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - |

| **Level of Effort Required** | | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
|---|---|---|---|---|---|---|---|---|---|---|
| ATLAS Computing Project Manager | Physicist | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE |
| Software Engineer | | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE |
| Associate Project Manager | Physicist | 0.20 FTE | 0.50 FTE | 0.50 FTE | 0.50 FTE | 0.50 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE |
| Administrative Support | | 0.50 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE | 1.00 FTE |
| Total FTEs | | 2.70 FTE | 3.50 FTE | 3.50 FTE | 3.50 FTE | 3.50 FTE | 4.00 FTE | 4.00 FTE | 4.00 FTE | 4.00 FTE |
| Offset (paid by others) | | 1.20 FTE | 1.50 FTE | 1.50 FTE | 1.50 FTE | 1.50 FTE | 2.00 FTE | 2.00 FTE | 2.00 FTE | 2.00 FTE |
| New FTEs Needed | | 1.50 FTE | 2.00 FTE | 2.00 FTE | 2.00 FTE | 2.00 FTE | 2.00 FTE | 2.00 FTE | 2.00 FTE | 2.00 FTE |

| **Computer Costs (FY99$)** | | October-98 | October-99 | October-00 | October-01 | October-02 | October-03 | October-04 | October-05 | October-06 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Computer Costs (FY99$)** | | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| Personnel Costs (FY99$) | | $ 148,611 | $ 148,611 | $ 148,611 | $ 222,916 | $ 525,217 | $ 525,217 | $ 876,424 | $ 876,424 | $ 876,424 |
| Hardware Costs (FY99$) | | $ 20,096 | $ 21,883 | $ 9,055 | $ 95,917 | $ 613,732 | $ 4,136,264 | $ 745,273 | $ 712,149 | $ 690,067 |
| Other Costs (FY99$) | | $ - | $ 12,000 | $ 30,000 | $ 30,000 | $ 60,000 | $ 60,000 | $ 60,000 | $ 60,000 | $ 60,000 |
| Total Costs (FY99$) | | $ 168,707 | $ 170,494 | $ 157,666 | $ 318,833 | $ 1,138,949 | $ 4,661,481 | $ 1,621,697 | $ 1,588,574 | $ 1,566,491 |
| | | | | | | | | | | |
| **CPU Power Required** | | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| CPU Required | | 100 Si95 | 100 Si95 | 100 Si95 | 500 Si95 | 5000 Si95 | 50000 Si95 | 50000 Si95 | 50000 Si95 | 50000 Si95 |
| CPU Purchased | | 100.00 Si95 | 25.00 Si95 | 25.00 Si95 | 425.00 Si95 | 4625.00 Si95 | 46250.00 Si95 | 12500.00 Si95 | 12500.00 Si95 | 12500.00 Si95 |
| CPU Retired | | 0.00 Si95 | 25.00 Si95 | 25.00 Si95 | 25.00 Si95 | 125.00 Si95 | 1250.00 Si95 | 12500.00 Si95 | 12500.00 Si95 | 12500.00 Si95 |
| **CPU Cost (FY99$)** | | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| CPU Moore's Factor | 1.50 | 1.00 | 0.67 | 0.44 | 0.30 | 0.20 | 0.13 | 0.09 | 0.09 | 0.09 |
| Linux CPU Price (FY99$/Si95) | | $ 168.28 | $ 112.19 | $ 74.79 | $ 49.86 | $ 33.24 | $ 22.16 | $ 14.77 | $ 14.77 | $ 14.77 |
| Sun CPU Price (FY99$/Si95) | **(Percent)** | $ 1,008.88 | $ 672.59 | $ 448.39 | $ 298.93 | $ 199.28 | $ 132.86 | $ 88.57 | $ 88.57 | $ 88.57 |
| Linux CPU Cost (FY99$) | 50% | $ 16,827.89 | $ 1,402.32 | $ 934.88 | $ 10,595.34 | $ 76,868.12 | $ 512,454.11 | $ 92,334.07 | $ 92,334.07 | $ 92,334.07 |
| Sun CPU Cost (FY99$) | 50% | $ - | $ 8,407.32 | $ 5,604.88 | $ 63,521.96 | $ 460,845.56 | $ 3,072,303.76 | $ 553,568.25 | $ 553,568.25 | $ 553,568.25 |
| *CPU Cost (FY99$)* | 100% | *$ 16,827.89* | *$ 9,809.64* | *$ 6,539.76* | *$ 74,117.29* | *$ 537,713.68* | *$ 3,584,757.87* | *$ 645,902.32* | *$ 645,902.32* | *$ 645,902.32* |
| **Disk Space Required** | | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| Disk Required | | 50 GB | 250 GB | 250 GB | 1000 GB | 5000 GB | 50000 GB | 50000 GB | 50000 GB | 50000 GB |
| Disk Purchased | | 50 GB | 200 GB | 63 GB | 813 GB | 4250 GB | 46250 GB | 12500 GB | 12500 GB | 12500 GB |
| Disk Retired | | 0 GB | 0 GB | 63 GB | 63 GB | 250 GB | 1250 GB | 12500 GB | 12500 GB | 12500 GB |
| **Disk Cost (FY99$)** | | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| Disk Moore's Factor | 1.50 | 1.00 | 0.67 | 0.44 | 0.30 | 0.20 | 0.13 | 0.09 | 0.06 | 0.04 |
| Disk Vault Price (FY99$/GB) | | $ 65.36 | $ 43.57 | $ 29.05 | $ 19.37 | $ 12.91 | $ 8.61 | $ 5.74 | $ 3.83 | $ 2.55 |
| Sun Raid Price (FY99$/GB) | | $ 208.33 | $ 138.89 | $ 92.59 | $ 61.73 | $ 41.15 | $ 27.43 | $ 18.29 | $ 12.19 | $ 8.13 |
| Sun SCSI Price (FY99$/GB) | **(Percent)** | $ 92.59 | $ 61.73 | $ 41.15 | $ 27.43 | $ 18.29 | $ 12.19 | $ 8.13 | $ 5.42 | $ 3.61 |
| Disk Vault Cost (FY99$) | 50% | $ 3,268.10 | $ 4,357.47 | $ 907.81 | $ 7,867.66 | $ 27,435.93 | $ 199,044.96 | $ 35,863.96 | $ 23,909.30 | $ 15,939.54 |
| Sun Raid Cost (FY99$) | 10% | $ - | $ 2,777.78 | $ 578.70 | $ 5,015.43 | $ 17,489.71 | $ 126,886.15 | $ 22,862.37 | $ 15,241.58 | $ 10,161.05 |
| Sun SCSI Cost (FY99$) | 40% | $ - | $ 4,938.27 | $ 1,028.81 | $ 8,916.32 | $ 31,092.82 | $ 225,575.37 | $ 40,644.21 | $ 27,096.14 | $ 18,064.09 |
| *Disk Cost (FY99$)* | 100% | *$ 3,268.10* | *$ 12,073.52* | *$ 2,515.32* | *$ 21,799.41* | *$ 76,018.46* | *$ 551,506.48* | *$ 99,370.54* | *$ 66,247.02* | *$ 44,164.68* |
| **Level of Effort Required** | | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| ARC Operation | | 0.33 FTE | 0.33 FTE | 0.33 FTE | 0.50 FTE | 1.00 FTE | 1.00 FTE | 2.50 FTE | 2.50 FTE | 2.50 FTE |
| ARC Administration | | 0.67 FTE | 0.67 FTE | 0.67 FTE | 1.00 FTE | 2.50 FTE | 2.50 FTE | 3.50 FTE | 3.50 FTE | 3.50 FTE |
| ARC Architecture Porting | | 0.00 FTE | 0.00 FTE | 0.00 FTE | 0.00 FTE | 0.00 FTE | 0.00 FTE | 0.00 FTE | 0.00 FTE | 0.00 FTE |
| *Level of Effort Required* | | *1.00 FTE* | *1.00 FTE* | *1.00 FTE* | *1.50 FTE* | *3.50 FTE* | *3.50 FTE* | *6.00 FTE* | *6.00 FTE* | *6.00 FTE* |
| **Fully burdened FTE Cost** | **Base Cost** | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| ARC Operation | $ 128,292 | $ 42,764 | $ 42,764 | $ 42,764 | $ 64,146 | $ 128,292 | $ 128,292 | $ 320,730 | $ 320,730 | $ 320,730 |
| ARC Administration | $ 158,770 | $ 105,847 | $ 105,847 | $ 105,847 | $ 158,770 | $ 396,925 | $ 396,925 | $ 555,695 | $ 555,695 | $ 555,695 |
| ARC Architecture Porting | $ 158,770 | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - | $ - |
| *Fully burdened FTE Cost* | | *$ 148,611* | *$ 148,611* | *$ 148,611* | *$ 222,916* | *$ 525,217* | *$ 525,217* | *$ 876,424* | *$ 876,424* | *$ 876,424* |
| TestBed Costs (FY99$) | | $ - | $ 10,000 | $ 25,000 | $ 25,000 | $ 50,000 | $ 50,000 | $ 50,000 | $ 50,000 | $ 50,000 |
| Software Costs (FY99$) | | $ - | $ 2,000 | $ 5,000 | $ 5,000 | $ 10,000 | $ 10,000 | $ 10,000 | $ 10,000 | $ 10,000 |
| | | | | | | | | | | |
| **Supporting Information** | | | | | | | | | | |
| 400 MHz PII Linux (128 MB/11 GB) | | $ 2,575 | 15.30 Si95 | | | | | | | |
| 14 node Sun E4000 (1GB) | | $ 250,000 | 247.80 Si95 | | | | | | | |
| 400 MHz PII Linux (256 MB/64 GB) | | $ 4,131 | 63 GB | | | | | | | |
| 400 MHz Quad Sun E450 (1GB) | | $ 30,000 | | | | | | | | |
| 8X18 GB Sun Raid (24 MB cache) | | $ 25,000 | 144 GB | | | | | | | |
| 9X18 GB Vanguard Disk Unit (Ultra Wide SCSI) | | $ 10,000 | 162 GB | | | | | | | |
| Sun/Sun Raid System | | $ 180,000 | 864 GB | | | | | | | |
| Sun/Vangard Disk System | | $ 90,000 | 972 GB | | | | | | | |

| Software Costs (FY99$) | | October-98 | October-99 | October-00 | October-01 | October-02 | October-03 | October-04 | October-05 | October-06 |
|---|---|---|---|---|---|---|---|---|---|---|
| | | **FY 1999** | **FY 2000** | **FY 2001** | **FY 2002** | **FY 2003** | **FY 2004** | **FY 2005** | **FY 2006** | **FY 2007** |
| | | | | | | | | | | |
| Pilot Project | | 3 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Control | | 2.5 | 4 | 4 | 4 | 3 | 3 | 2 | 2 | 1 |
| Database | | 1 | 2 | 4 | 4 | 6 | 6 | 6 | 6 | 4 |
| Collaboration Design Support | | 0 | 1 | 2 | 2 | 3 | 3 | 3 | 2 | 2 |
| MONARC | | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Contingency | | 0 | 0 | 1 | 2 | 2 | 2 | 2 | 2 | 1 |
| Total Computer Professi | | 7.5 | 10 | 11 | 12 | 14 | 14 | 13 | 12 | 9 |
| Offset (Paid by others) | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Computing Professionals Needed | | 7.5 | 10.0 | 11.0 | 12.0 | 14.0 | 14.0 | 13.0 | 12.0 | 9.0 |
| | | | | | | | | | | |
| **Software Costs (FY99$)** | | | | | | | | | | |
| Fully Burdened Salary | $175,000 | $1,312,500 | $1,750,000 | $1,925,000 | $2,100,000 | $2,450,000 | $2,450,000 | $2,275,000 | $2,100,000 | $1,575,000 |
| CP Travel | $13,000 | $97,500 | $130,000 | $143,000 | $156,000 | $182,000 | $182,000 | $169,000 | $156,000 | $117,000 |
| Workstation Cost | $3,000 | $22,500 | $30,000 | $33,000 | $36,000 | $42,000 | $42,000 | $39,000 | $36,000 | $27,000 |
| Total SWF cost | | $1,432,500 | $1,910,000 | $2,101,000 | $2,292,000 | $2,674,000 | $2,674,000 | $2,483,000 | $2,292,000 | $1,719,000 |

# ATLAS Cost Summary (FY99$)

| | October-98 | October-99 | October-00 | October-01 | October-02 | October-03 | October-04 | October-05 | October-06 |
|---|---|---|---|---|---|---|---|---|---|
| | FY 1999 | FY 2000 | FY 2001 | FY 2002 | FY 2003 | FY 2004 | FY 2005 | FY 2006 | FY 2007 |
| **FTE Summary** | | | | | | | | | |
| **ATLAS Computing Management** | | | | | | | | | |
| Total FTEs Needed | 2.7 FTE | 3.5 FTE | 3.5 FTE | 3.5 FTE | 3.5 FTE | 4.0 FTE | 4.0 FTE | 4.0 FTE | 4.0 FTE |
| Offset (funded by others) | 1.2 FTE | 1.5 FTE | 1.5 FTE | 1.5 FTE | 1.5 FTE | 2.0 FTE | 2.0 FTE | 2.0 FTE | 2.0 FTE |
| New FTEs Needed | 1.5 FTE | 2.0 FTE | 2.0 FTE | 2.0 FTE | 2.0 FTE | 2.0 FTE | 2.0 FTE | 2.0 FTE | 2.0 FTE |
| **ATLAS Regional Center** | | | | | | | | | |
| Total FTEs Needed | 2.5 FTE | 3.9 FTE | 5.4 FTE | 6.4 FTE | 11.0 FTE | 11.0 FTE | 14.5 FTE | 14.5 FTE | 14.5 FTE |
| **Software Development** | | | | | | | | | |
| Total FTEs Needed | 7.5 FTE | 10.0 FTE | 11.0 FTE | 12.0 FTE | 14.0 FTE | 14.0 FTE | 13.0 FTE | 12.0 FTE | 9.0 FTE |
| Offset (funded by others) | 0.0 FTE | 0.0 FTE | 0.0 FTE | 0.0 FTE | 0.0 FTE | 0.0 FTE | 0.0 FTE | 0.0 FTE | 0.0 FTE |
| New FTEs Needed | 7.5 FTE | 10.0 FTE | 11.0 FTE | 12.0 FTE | 14.0 FTE | 14.0 FTE | 13.0 FTE | 12.0 FTE | 9.0 FTE |
| **Cost Summary** | | | | | | | | | |
| **ATLAS Computing Management** | | | | | | | | | |
| Personnel Costs (FY99$) less offset | $213,000 | $256,000 | $256,000 | $256,000 | $256,000 | $256,000 | $256,000 | $256,000 | $256,000 |
| Hardware Costs (FY99$) | $0 | $0 | $0 | $0 | $0 | $0 | $0 | $0 | $0 |
| Other Costs (FY99$) | $28,000 | $24,000 | $24,000 | $24,000 | $24,000 | $24,000 | $24,000 | $24,000 | $24,000 |
| Total Funding Needed (FY99$) | $241,000 | $280,000 | $280,000 | $280,000 | $280,000 | $280,000 | $280,000 | $280,000 | $280,000 |
| **ATLAS Regional Center** | | | | | | | | | |
| Personnel Costs (FY99$) less offset | $427,876 | $683,042 | $1,213,456 | $1,421,700 | $2,364,822 | $2,364,822 | $3,024,426 | $3,024,426 | $3,024,426 |
| Hardware Costs (FY99$) | $20,096 | $21,883 | $13,583 | $143,875 | $1,511,668 | $6,761,788 | $1,826,449 | $1,494,313 | $1,436,999 |
| Other Costs (FY99$) | $18,000 | $96,500 | $150,750 | $187,000 | $223,000 | $223,000 | $162,500 | $162,500 | $162,500 |
| Total Funding Needed (FY99$) | $465,972 | $801,425 | $1,377,789 | $1,752,575 | $4,099,490 | $9,349,609 | $5,013,375 | $4,681,239 | $4,623,925 |
| **Software Development** | | | | | | | | | |
| Personnel Costs (FY99$) | $1,312,500 | $1,750,000 | $1,925,000 | $2,100,000 | $2,450,000 | $2,450,000 | $2,275,000 | $2,100,000 | $1,575,000 |
| Workstations (FY99$) | $22,500 | $30,000 | $33,000 | $36,000 | $42,000 | $42,000 | $39,000 | $36,000 | $27,000 |
| Other Costs (FY99$) | $97,500 | $130,000 | $143,000 | $156,000 | $182,000 | $182,000 | $169,000 | $156,000 | $117,000 |
| Total Funding Needed (FY99$) | $1,432,500 | $1,910,000 | $2,101,000 | $2,292,000 | $2,674,000 | $2,674,000 | $2,483,000 | $2,292,000 | $1,719,000 |
| TOTAL COST | $2,139,472 | $2,991,425 | $3,758,789 | $4,324,575 | $7,053,490 | $12,303,609 | $7,776,375 | $7,253,239 | $6,622,925 |